

تطبيقاتٌ أساسيَّة في المُعالَجة الآليَّة للُّغة العربيَّة

مباحث لغوية ٥٨

تحرير

د. مُحسن رَشوان　　د. المُعتزّ بالله السَّعيد

الباحثون:

د. مُحمَّد عَطيَّة　　د. مُحمَّد عفيفي

د.شريف مهدي عَبده　　د. علي علي فهمي

مباحث لغوية ٥٨

تطبيقاتٌ أساسيَّة في المُعالَجة الآليَّة للُّغة العربيَّة

تحرير

د. مُحسن رَشوان د. المُعتزّ بالله السَّعيد

الباحثون:

د. مُحمَّد عَطيَّة د. مُحمَّد عفيفي

د.شريف مهدي عَبده د. علي علي فهمي

١٤٤١هـ - ٢٠١٩م





تطبيقاتٌ أساسيَّة في المُعالَجة الآليَّة للُّغة العربيَّة

الطبعة الأولى
١٤٤١ هـ - ٢٠١٩ م

المملكة العربية السعودية - الرياض
ص.ب ١٢٥٠٠ الرياض ١١٤٧٣
هاتف:٠٠٩٦٦١١٢٥٨٧٢٦٨ - ٠٠٩٦٦١١٢٥٨١٠٨٢
البريد الإليكتروني: nashr@kaica.org.sa

مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية، ١٤٤١هـ.
فهرسة مكتبة الملك فهد الوطنية أثناء النشر
رشوان، محسن
تطبيقات أساسية في المعالجة الآلية للغة العربية./ محسن رشوان؛ المعتز بالله السعيد .-الرياض، ١٤٤٠هـ
..ص؛ .. سم
ردمك: ٢ -٥٣- ٨٢٢١- ٦٠٣- ٩٧٨
١ - اللغة العربية - معالجة البيانات أ. السعيد ، المعتز بالله (مؤلف مشارك) ب. العنوان
ديوي ٤١٠,٢٨٥ ١٤٤٠/١٠١٦٧
رقم الإيداع: ١٤٤٠/١٠١٦٧
ردمك: ٢ -٥٣- ٨٢٢١- ٦٠٣- ٩٧٨

التصميم والإخراج
دار وجوه للنشر والتوزيع
Wojooh Publishing & Distribution House
WOJOOOH
www.wojoooh.com
المملكة العربية السعودية - الرياض
الهاتف:4562410 الفاكس:4561675
للتواصل والنشر:
info@wojoooh.com



فهرس الكتاب

كلمة المركز

يعمل المركز في مجال البحث العلمي ونشر الكتب مستهدفاً التركيز على المجالات البحثية التي ما زالت بحاجة إلى تسليط الضوء عليها، وتكثيف البحث فيها، ولفت أنظار الباحثين والجهات الأكاديمية إلى أهمية استثمارها بمختلف وجوه الاستثمار، وذلك مثل مجال (التخطيط اللغوي) و (العربية في العالم) و(الأدلة والمعلومات) و (تعليم العربية لأبنائها أو لغير الناطقين بها) إلى غير ذلك من المجالات، وإن من أهم مجالات البحث المستقبلية في اللغة العربية مجال (العربية والحوسبة ، والذكاء الاصطناعي) حيث إن حياة اللغات ومستقبلها مرهونة بمدى تجاوبها مع التطورات التقنية والعالم الافتراضي، وكثافة المحتوى الالكتروني المكتوب، وهو ما يشكّل تحديا حقيقيا أمام اللغات غير المنتجة للمعرفة أو للتقنية.

وقد عمل المركز على تسليط الضوء على هذا المجال التخصصي؛ مستعينا بالكفاءات القادرة من المهتمين بالتخصص البيني (بين اللغة والحاسوب) مقدّرا جهودهم، وهادفاً إلى نشرها، وتعميم مبادئها، راغباً أن يكون هذا المسار العلمي مقررا في الجامعات في كلية العربية والحاسوب، ومجالا بحثيا يقصده الباحثون الأكديميون، والجهات البحثية العربية.

وقد أصدر المركز سابقا ستة عشر كتاباً مختصا في (حوسبة العربية) وفي الإفادة من (المدونات اللغوية) في الأبحاث العربية، ويحتفل بإصدار سبعة كتب جديدة مختصة في (حوسبة العربية والذكاء الاصطناعي) ، ويقدمها للقارئ العربي، وللجهات الأكاديمية؛ للإفادة منها في مناهج التعليم والبناء عليه، وهذه الكتب السبعة هي: (العربيَّة والذَّكاء الاصطناعيّ، تطبيقات الذكاء الاصطناعي في خدمة اللغة العربية، خوارزميات الذكاء الاصطناعي في تحليل النص العربي، مقدمة في حوسبة اللغة العربية، الموارد اللغوية الحاسوبية، المعالجة الآلية للنصوص العربية، تطبيقات أساسية في المعالجة الآلية للغة العربية).

ويشكر المركز السادة مؤلفي الكتب، ومحرريها، لما تفضلوا به من عمل علمي رصين، وأدعو الباحثين والمؤلفين إلى التواصل مع المركز لاستكمال المسيرة، وتفتيق فضاءات المعرفة.

وفق الله الجهود وسدد الرؤى.

الأمين العام

أ. د. محمود إسماعيل صالح

مقدمة

تشهدُ مُعالَجَةُ اللُّغات الطَّبيعيَّة ظُهورَ أساليبَ مُبتَكَرة؛ ينطلقُ بعضُها من منطق اللُّغة القائم على استيعاب قوانين هذه اللُّغات وقواعدِها؛ وينطلقُ بعضُها الآخَرُ من منطق الآلة القائم على نمذجة اللُّغة في مُستوياتها الـمُتعدِّدة. والواقعُ أنَّ الأساليبَ والمناهج المستخدمة في مُعالجة اللُّغات الطَّبيعيَّة تتَّسمُ بقدرٍ من المرونة الَّتي تسمحُ بالجمع بينَ مُعطياتٍ لُغويَّةٍ وإحصائيَّة، على النَّحو الَّذي يُمَكِّنُ من الوصول إلى أفضل النَّتائج الـمُمكنة في تطبيقات حوسبة اللُّغات الطَّبيعيَّة.

إنَّ الوقوفَ على نجاعة هذه الأساليب وصلاحيَّتِها للـمُعالَجَة يستدعي تقييمًا موضوعيًّا، يتجاوزُ الإطارَ النَّظريَّ الافتراضيَّ إلى إطارٍ تطبيقيٍّ عمليّ، يُساعدُ على استكشاف إشكالات الـمُعالَجَة وعَقَباتِها، ويُقدِّمُ حلولًا فعَّالةً لتجاوُزها. ولعلَّ مثلَ ذلك التَّقييم يضعُ أيدينا على حقيقةٍ، مفادُها أنَّ أساليبَ مُعالجة اللُّغات الطَّبيعيَّة تتكاملُ فيما بينها؛ فيُعَوَّلُ على بعضِها في بناء تطبيقاتٍ مُعَيَّنة، ويُعوَّلُ على أخرى في تطوير هذه التَّطبيقات.

من هذا الـمُنطَلَق، نُقدِّمُ للقارئ العربيّ الكتابَ الرَّابعَ من سلسلة دراسات وبُحُوث في حَوسَبَة اللُّغة العربيَّة، بعُنوان (تطبيقات أساسيَّة في الـمُعالَجَة الآليَّة للُّغة العربيَّة)؛ ونُحاولُ من خلالِهِ أن نطرحَ رُؤيةً ذاتَ بُعدٍ تطبيقيّ حولَ تقنيات مُعالَجَة اللُّغة العربيَّة

في صُورَتَيها: المكتوبة والمنطوقة، ونعرُجُ من ذلكَ أيضًا إلى جوانب الإفادة من تطبيقات حَوسَبة اللُّغة العربيَّة في تعليمها من ناحية، وتقييم مُخرَجات عمليَّة التَّعليم من ناحيةٍ أخرى.

ورغبةً في تحقيق أهدافِنا المنشودة، فقد قسَّمنا الكتابَ إلى أربعة فُصُولٍ، على النَّحو الآتي:

- الفصل الأوَّل: مُعالَجة النَّصّ العربيّ المكتوب؛ يُعنى بتقنيات التَّعرُّف على النَّصّ العربيّ المكتوب، ويشتملُ على ثلاثة مباحث؛ حيثُ يتناولُ في المبحث الأوَّل طبيعةَ رسم النَّصّ العربيّ، ويُقدِّمُ لذلك بإرهاصات الخِطاطة العربيَّة وتاريخ حوسَبَتِها وتحدِّيات التَّعرُّف عليها آليًّا؛ ويُعنى المبحثُ الثَّاني بأنواع التَّعرُّف الآليّ على النَّصّ العربيّ المكتوب وتطبيقاته. أمَّا المبحثُ الثَّالثُ فيعرِضُ لأساليب التَّعرُّف على النَّصّ العربيّ، سواءٌ أكانَ مطبوعًا أم مخطوطًا؛ ويعرِضُ كذلكَ للموارد اللُّغويَّة اللَّازمة لتدريب وتقويم أنظمة التَّعرُّف على النَّصّ العربيّ.
- الفصل الثاني: مُعالَجة النَّصّ العربيّ المنطوق؛ يُعنى بتقنيات التَّعرُّف على الكلام المنطوق؛ ويشتملُ على ثلاثة مباحثَ؛ حيثُ يُقدِّمُ المبحثُ الأوَّلُ تمهيدًا حولَ التَّعرُّف الآليّ على الكلام ومُكوِّنات نُظُمِه، مع العناية بتطبيق ذلكَ في اللُّغة العربيَّة. ويُعنى المبحثُ الثَّاني بنُظُم تحويل النَّصّ (المكتوب) إلى كلامٍ (منطوق). ويعرِضُ المبحثُ الثَّالثُ لنُظُم التَّعرُّف على اللُّغة والـمُتكلِّم.
- الفصل الثَّالث: تطبيقات مُعالَجة اللُّغة العربيَّة في مجال التَّعليم؛ وترتكزُ مادَّةُ هذا الفصل على آليَّات توظيف الحاسوب في تعليم اللُّغات؛ حيثُ يُقدِّمُ تمهيدًا حولَ تقنيات مُعالَجة اللُّغات الطَّبيعيَّة وجدواها في تعلُّم اللُّغة؛ ويُقدِّمُ الفصلُ تطبيقًا عمليًّا حولَ تعلُّم النُّطق باستخدام تقنية التَّعرُّف على الكلام المنطوق، وتطبيقًا آخرَ حولَ تعلُّم الكتابة باستخدام تقنية التَّعرُّف على النَّصّ المكتوب.
- الفصل الرَّابع: التَّقييم الآليّ؛ يتناولُهُ الفصلُ باعتبارِهِ أحدَ تطبيقات التَّنقيب في النُّصُوص. ويُقدِّمُ تمهيدًا حولَ الأسئلة المقاليَّة وأنواعها، ثُمَّ يعرِضُ لطُرُق التَّقييم الآليّ للمقال، وطُرُق تقييم الإجابات القصيرة، وأنظمة تقييم

الرِّياضيَّات، وأنظمة الكشف عن السَّرقات الأدبيَّة. كما يعرضُ لواقع أنظمة التَّقييم الآليّ في العربيَّة.

وبعدُ؛ فإنَّنا نرجو أن يُساعدَ هذا الكتابُ على فهمٍ أعمقَ لتقنيات مُعالجة اللُّغة العربيَّة وتطبيقاتها الأساسيَّة؛ ونلتمسُ أن تكونَ مادَّةُ هذا الكتابِ مفتاحًا للباحثينَ في ميادين حوسبة اللُّغة للبحث عن وسائلَ مُبتكَرة لبناء وتطوير تقنيات مُعالجَة العربيَّة عبرَ مُستوياتها الـمُختلفة.

نسألُ الله تعالى أن يتقبَّلَ هذا الجهدَ بالذِّكر الحَسَنِ والأجرِ الجزيل، وأن يجعلَه من العلمِ الَّذي ينفعُ أصحابَه بعد مماتهم.

ربَّنا عليكَ توكَّلنا وإليكَ أنَبنا وإليكَ المصِير.

الـمُحرِّران

الفصل الأول

مُعالجَة النَّصّ العربيّ المكتوب

د. مُحمَّد عطيَّة

المبحث الأوَّل: طبيعة رسم النَّصّ العربيّ (الخِطاطة).
المبحث الثَّانِي: أنواع التَّعرُّف الآليّ على النَّصّ العَرَبيّ المكتوب.
المبحث الثَّالث: أساليب التَّعرُّف الآليّ على النَّصّ العربيّ المكتوب.

المبحث الأوَّل

طبيعة رسم النص العربي

(الخِطَاطَة)

١ – خلفية تاريخية لتطور الخِطاطة العربية.

٢- تاريخ حوسبة الخِطاطة العربية.

٣- تحديات الخِطاطة العربية التي تواجه التعرف عليها آليًّا.

١- خلفية تاريخية لتطور الخِطاطة العربية

حافظت اللغة العربية على استمراريتها كلغة محكية محتفظة بكيانها منذ ما يربو على ألفي عام - على أقل تقدير- إلى وقتنا هذا، مما يتوجها دون شك كأعرق اللغات الحية الكبرى المتداولة في العالم المعاصر. وقد احتفظت هذه اللغة على مدى هذا التاريخ الطويل بقدرتها على الاستجابة لمختلف الاحتياجات الحضارية للأجيال المتتابعة، كما استجابت باقتدار للتوسعات الجغرافية والتمددات العرقية للمتحدثين بها الذين تزايدوا من عدد محدود محصور أساسًا في شبه الجزيرة العربية قبيل ظهور الإسلام حتى تجاوز عددهم عام ٢٠١١م الثلاثمئة وستين مليونًا يتخذون من العربية لغتهم الأم فضلاً عن عشرات الملايين الآخرين الذين يتحدثونها ضمن ما يربو على مليار مسلم غير عربي.

ورغم أن قواعد الصوتيات والصرف والنحو ... وغير ذلك من أركان اللغة ودعائمها قد احتفظت بسماتها المميزة عبر ذلك التاريخ الممتد والتوسع المستمر، فإن هناك تطورات متتابعة كانت تجري على بعض الملامح الأخرى لِلُّغة. وتأتي الخِطاطة العربية (التي تحدد طبيعة رسم النص العربي) ضمن الملامح التي جرت عليها مثل هذه التطورات وهي نفس الخِطاطة المستخدمة لرسم النصوص في عدد آخر من اللغات الشرقية المعاصرة كالأُرْدِيَّة والفارسية والكردية، كما استخدمت لرسم اللغة التركية حتى ثلاثينات القرن العشرين الميلادي.

لم يكن الغالب على سكان شبه الجزيرة العربية قبل الإسلام تكوين دول ذات أنظمة مستقرة ومعاملات معقدة، وإنما كانوا في الغالب قبائل بدوية، ولم تكن الكتابة شائعة بينهم - حتى أنهم عرفوا وقتذاك بالأميين - حيث اعتمدوا أساسًا في تداول المعلومات والمعرفة على رواية وحفظ الأقوال البليغة المختصرة من شعر ونثر، وهو ما تميزوا فيه واشتهروا به. أما القلة النادرة التي كانت تعرف الكتابة - ربما بسبب مزاولتها للتجارة أو لوجودها على تخوم الجزيرة واحتكاكها بكبرى الحضارات المعاصرة آنذاك - فقد اكتفت بنظام كتابة يميل إلى البساطة.

كانت الخطاطة في ذلك الوقت الباكر تمثل الحروف العربية الثمانية والعشرين (أو الثلاثين) بخمسة عشر أو ستة عشر رمزًا رسوميًّا «جرافيم» (grapheme) ومن

الواضح أن ذلك يسبب التباسًا كبيرًا في تمييز الحروف. وعلى سبيل المثال فإن النقاط لم تكن قد أدخِلت بعد في الخطاطة العربية، ولذلك فقد كانت حروف الحاء والخاء والجيم كلها تُمَثَّل بنفس الرمز الرسومي، وكذلك كانت حروف الباء والتاء والثاء والنون والياء كلها تمثل بنفس الرمز الرسومي، ...، إلخ. وقد تمكن مَنْ يكتبون العربية في ذلك الزمن من التعايش مع هذا الالتباس بسبب ملكاتهم اللغوية الرفيعة، وربما كذلك بسبب البساطة النسبية للرسائل المتداولة بينهم.

ويبين الشكل التالي كيف كانت ستبدو العبارة الآتية «الترجمة وسيلة أساسية لتبادل الحضارات بين الشعوب على مر العصور» عند خِطاطتها في تلك المرحلة الباكرة.

الىرحمه وسىله اساسىه لىىادل الحصاراٮ ٮىں السعوٮ على مر العصور

الشَّكل ١-١: مثال على عبارة مكتوبة بالخطاطة العربية القديمة ما قبل الإسلام.

وفي بدايات القرن السابع الميلادي بُعِثَ الرسول الكريم محمد - صلى الله عليه وسلم - برسالة الإسلام وفي القلب منها القرآن الكريم بلسان عربي مبين وقد أمر الرسول الكريم أصحابه بتدوينه أولاً بأول وكان ذلك عبر الخِطاطة القديمة، ثم جمعه من بعد ذلك خلفاؤه الراشدون ونشروا نُسَخَهُ في البلاد التي وصل إليها الإسلام في زمنهم والتي غطت مساحات شاسعة من الأرض تسكنها أجناس شتى بألسنة شتى دخلوا في دين الإسلام ولما يتقنوا اللغة العربية بعد (وهم من اصطُلِح على تسميتهم بالأعاجم). وبطبيعة الحال فقد كثرت وتواترت أخطاء هؤلاء المسلمون الأعاجم في قراءة القرآن (وهو ما اصطُلِحَ على تسميته باللَّحْن) وكانت الخِطاطة القديمة للعربية من أسباب اللَّحْن لديهم حيث كانت تغيب الحركات الصوتية وتلتبس الحروف المتشابهة فيخلطون بينها. ولما انتشر ذلك خاف العلماء على كتاب الإسلام من التحريف فأضيفت النقاط للتمييز بين الحروف المتشابهة لتمييزها بصورة حاسمة كما نعرفها اليوم، وينسب ذلك العمل إلى «نصر بن عاصم الليثي»، ويبين الشكل رقم (١-٢) أدناه نفس العبارة التي أوردناها كمثال في شكل رقم (١-١) أعلاه لكنها مخطوطة بنقط الحروف.

الترجمة وسيلة أساسية لتبادل الحضارات بين الشعوب على مر العصور

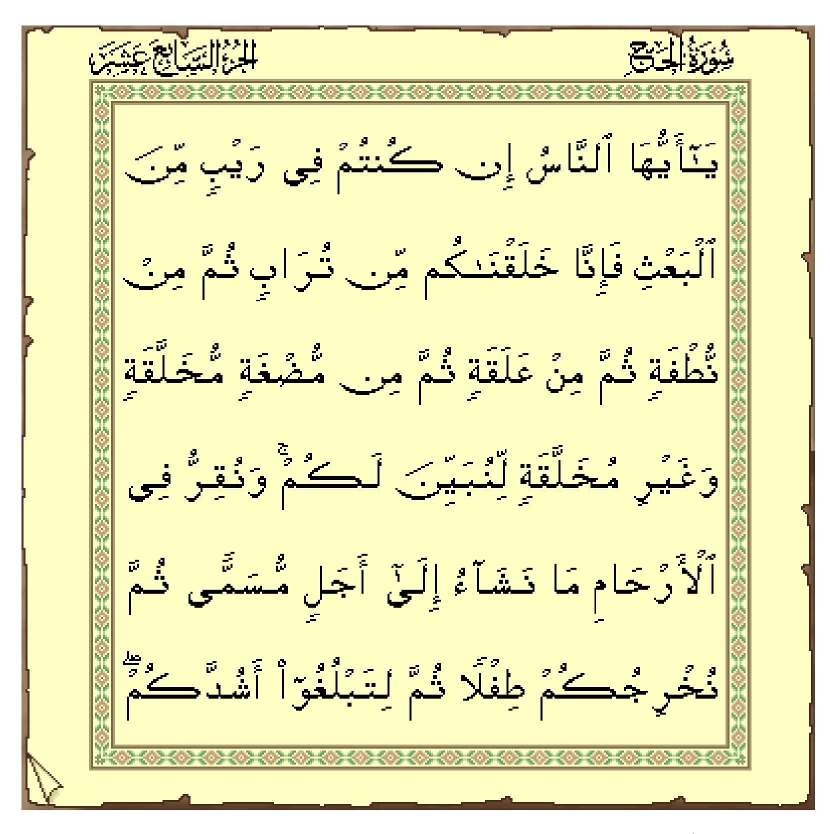
الشَّكل ١-٢: نفس العبارة في الشكل السابق، ولكنها مخطوطة بالنقاط لتقليل الالتباس.

وفي حين أن نقاط «نصر بن عاصم» قد حسمت الالتباس بين رسم حروف الهجاء، فقد بقي اللحن نتيجة الخلط بين الحركات الصوتية لكلٍّ من هذه الحروف، وذلك ما أزاله اللغوي الشهير «أبو الأسود الدُّؤَلِيّ» بحلٍّ ناجع يتمثل في وضع نقاط إضافية فوق أو تحت كل حرف للدلالة على الحركة الصوتية المصاحبة له وسمى تلك النقاط «نقاط الإعراب» (والإعراب هنا لا يقصد به ما نصطلح عليه اليوم من إبانة موقع الكلمة النحوي، ولكنه قصد به آنذاك إبانة الحركة الصوتية المصاحبة لكل حرف).

وبعد ذلك بعقود قليلة قام أستاذ أساتذة اللغة العربية في عصره «الخليل بن أحمد الفراهيدي البصري»[1] بتحسين رسم «نقاط الإعراب» التي وضعها «أبو الأسود الدؤلي» وأبدلها بعلامات التشكيل (أو الضبط) الصوتي التي نعرفها اليوم من فتحة وكسرة وضمة وسكون، فضلاً عن أنه أضاف على عمل «نصر بن عاصم» علامةً تدل على تضعيف الحرف (الشَّدَّة) ورَسْمَ الهمزة.

وجعلت الخِطاطة العربية بعد إضافات «الخليل ابن أحمد» النص العربي المكتوب -وخصوصًا القرآن الكريم- يسير القراءة واستبعدت إلى حد بعيد حالات الالتباس فيه. واستمر تلامذة الخليل في إضافة تحسينات تكميلية (كعلامات الوقف، والوصل، والمد، والإدغام ... إلخ) على رسم القرآن الكريم حتى بلغ ذروته في كمال الضبط بنهايات القرن الثالث الهجري حين بدا المصحف الشريف على الهيئة التي نألفها اليوم (انظر شكل رقم ١-٣ أدناه) حيث تشتمل خِطاطة النص القرآني في المصحف (وهي ما نعرفها اليوم بالرسم العثماني) على كل الرموز اللازمة لتمكين القارئ من تلاوة القرآن الكريم كما أنزل على الرسول الكريم محمد صلى الله عليه وسلم.

١- جدير بالذكر أن «الخليل» قد تتلمذ على يديه كبار المشاهير من أساطين العربية ومنهم «سِيبَوَيْه» و«الأَصْمَعِيّ» و«الكِسائيّ» و«النَّضْر بن شُمَيْل» ... وغيرهم.



الشَّكل ١-٣: عينة من نص القرآن الكريم مخطوطة بالرسم العثماني.

ومع صعود الحضارة الإسلامية وقيام دولها العظمى المتعاقبة من العباسيين إلى العثمانيين مرورًا بما بينهما، فإن النشاط الإداري والسياسي وكذلك النهضة العلمية والثقافية جعلت جميعها من التدوين والتوثيق نشاطًا وفنًّا كبيرًا قائمًا بذاته، وبالتالي فإن الخطاطة العربية قد نالت حظًّا عظيمًا من الاهتمام وظهر مفهوم «الخطوط» (الفُونْتات/ الأبناط fonts) حيث يمثل كل «خَطّ» نسقًا مطردًا يلتزم بقواعد الخِطاطة العربية العامة المتفَق عليها، ويتميز في الوقت نفسه بجماليات خاصة به تفيد أغراضًا معينة سواء كانت الزخرفة والتشكيل (مثل الثُّلُث والديواني؛ انظر الشكل رقم ١-٤ أدناه)، أو الإيحاء بالعراقة والأصالة (مثل الكوفي)، أو الرشاقة والوضوح والمعيارية (عائلة خطوط النسخ؛ انظر الشكل ١-٥ أدناه)، أو السهولة والسرعة (خط الرقعة؛ انظر الشكل رقم ١-٦ أدناه) ... إلخ.



الشَّكل ١-٤: عينة زخرفية من الخط الديواني.

خطوط النسخ هي أكثر الخطوط العربية شيوعًا في المطبوعات
وهي تجمع بين سهولة القراءة وجمال الشكل وقابلية الطباعة
وتضم هذه العائلة من الخطوط العديد من الأفراد كما في هذا المثال

الشَّكل ١-٥: عينة من خطوط النسخ.

يعتبر خط الرقعة من الخطوط الحديثة نسبياً لأن
الاهتمام به بزغ في النصف الثاني من القرن الثالث الهجري
فوضع له الخطاطون قواعد ضبطٍ وأسس ثابتة
وللخطاطين العثمانيين اليد
الطولى في إرساء قواعده الأخيرة
وخاصة الخطاط محمد عزت

الشَّكل ١-٦: عينة من خط الرقعة.

٢- تاريخ حوسبة الخِطاطة العربية

يمكن النظر لحوسبة خِطاطة اللغة بصفة عامة على أنها امتداد لسالفاتها من التقنيات الأقدم في هذا الصدد وهي الآلات الكاتبة الكهربائية ومن قبلها الميكانيكية ومن قبلهما تقنيات الطباعة بأجيالها المتعاقبة. ومنذ بداية رحلة العالم مع الطباعة لم تكن اللغة العربية في هذا الصدد متخلفة عن نظيراتها من اللغات الأوروبية، فقد أدخل العثمانيون آليات الطباعة مبكرًا إلى حاضرتهم المركزية «الآستانة» في النصف الثاني من القرن الخامس عشر، وهناك خلاف عما إذا كانوا سبقوا إليها بشكل مستقل أو نقلوها عن الأوروبيين الألمان (حيث استطاع «جوتِنْبِرْج» إنشاء أول مطبعة في «مِينْز» عام ١٤٤٨م)، وبغض النظر عن ذلك فإن الأمر المهم هو أن العثمانيين وقتذاك كانوا يخطون اللغة التركية (لغة دولتهم المركزية الرسمية) بالحرف العربي (واستمر ذلك حتى ثلاثينات القرن العشرين الميلادي) كما أن اللغة العربية نفسها كانت لغة سائدة في العلم والثقافة يتقنها جُلُّ المتعلمين الأتراك.

ودخلت عملية الطباعة إلى العالم العربي مع بدايات القرن التاسع عشر الميلادي في مصر وبلاد الشام، وأخذت من ذلك الوقت في الانتشار والتوسع؛ ومن مظاهر ذلك مثلاً إنشاء «محمد علي باشا الكبير» للمطابع الأميرية في مصر، ثم تواصلت مواكبة نسخ وطباعة الخط العربي بعد ذلك لكل تطور في هذا المجال حتى وصلنا لعصر الحاسبات الرقمية في النصف الثاني من القرن العشرين الميلادي حيث لم تكن الخِطاطة العربية أقل استعدادًا من أية لغة كبرى أخرى للرقمنة والحوسبة. وعلى الرغم من ذلك فقد تأخرت حوسبة الخِطاطة العربية بشكل مستقر ومقبول إلى نهاية ثمانينات وبداية تسعينات القرن العشرين الميلادي! ونرى أنه من الهام إفساح المساحة في هذا المقام لإزالة بعض الأوهام والالتباسات التي شاعت بأن هذا التأخر كان بسبب قصور وتعقيد في الخِطاطة العربية أو بسبب تقصير وعجز علماء ومهندسي الحوسبة العرب.

ظهرت بواكير الحاسبات الإلكترونية الرقمية في أربعينات وخمسينات القرن العشرين الميلادي في الولايات المتحدة الأمريكية، وكانت وقتئذٍ أجهزة باهظة التكاليف هائلة الحجم (قد يشغل الواحد منها مبنى بأكمله) معقدة التشغيل وفي بعض الأحيان سرية التفاصيل، ومع ذلك فإن أرفعها أداءً كان ذا إمكانات حسابية أقل من الآلات الحاسبة في أيامنا هذه! وكان تفاعل هذه الأجهزة مع مشغليها من المهندسين المتخصصين محدودًا وكان من البديهي أن يكون هذا التفاعل عبر اللغة الإنجليزية. وفي العقدين التاليين أخذت إمكانات واعتمادية هذه الأجهزة في التصاعد بشكل ملحوظ (وإن ظل أقواها أضعف كثيرًا من أقل حاسب شخصي في أيامنا هذه) بينما أخذ الطلب على تطبيقاتها في مجالات الإدارة والهندسة والتنظيم فضلاً عن البحث العلمي في البروز والتنامي، ومع ذلك ظلت ضخمة الحجم (يشغل الواحد منها قاعة كبيرة ويستخدم طاقة كهربية هائلة) باهظة التكلفة، وعلى ذلك شرعت شركات - مثل «آي بي إم» الأمريكية - في تصنيع أعداد محدودة منها لبيعها بشكل تجاري إلى كبريات المؤسسات المصرفية والصناعية والحكومية والبحثية، ولما كان تشغيلها ما يزال يحتاج إلى خبراء ذوي تأهيل هندسي متخصص - وهم يتقنون الإنجليزية بطبيعة الحال - فإن انفراد اللغة الإنجليزية بواجهات استخدام برمجيات تلك الحاسبات ظل أمرًا مقبولاً بعيدًا عن الانتقاد والجدل.

تغيــر المشهد جذريًّا بعد تصنيع وشيوع المشغلات الدقيقة للحاسبات «المايِكْرُو-بروسِيسُور» (Microprocessor) على نطاقٍ واسعٍ مطلَعَ الثمانينات حيث شاعت حاسباتٌ صغيرة للألعاب وللتطبيقات البسيطــة بين عمـــوم الناس وخاصة الشباب (ومن أشهرها في ذلك الوقت تلك التي أطلقتها شركة «أَتَارِي»، وشركة «سِينْكِلِير»)، ورغم أن تلك الحاسبات التي ارتكزت على الأجيال الأولى من المشغــلات الدقيقــة كانـــت ذات إمكانات بدائية جدًّا مقــارنة بالحاسبـــات الشخصية في أيامـــنا هـــذه فإنها لاقـــت رواجـــًا هائلاً بين الشباب في ذلك الوقـــت حـــيث كانـــت مدخلهم إلى عالم رقمي جديد ساحر باهر. وعند هذه النقطة بدأ الالتفات إلى مسألة اللغة التي تتفاعل بها برمجيات تلك الحاسبات مع مستخدميها الذين لم يكن جميعهم بالضرورة يتقنون الإنجليزية، ونشأ عندئذٍ الطلب على استخدام اللغات المحلية للأسواق التي تنتشر بها أعداد كبيرة من مستخدمي تلك الحاسبات وكانت العربية من بين تلك اللغات.

ما عقَّد وأخَّر تلبية هذا الطلب هو الأسباب الثلاثة المترابطة الآتية:

١. المحدودية الشديدة لقدرات ذاكرة تلك الحاسبات وقتئذٍ (تراوحت بين ثمانية آلاف كلمة إلى خمسة وستين ألف كلمة كحد أقصى) مما يعني أن كل البرمجيات العاملة وبياناتها في آن واحد لا يمكن أن يتعدى حجمها هذه الذاكرة الضئيلة. وكان حيز الكلمات الرقمية التي تستخدم لتمثيل عناصر المعلومات - ومنها الحروف - ضيقةً جدًّا؛ بدأت بأربعة أرقام ثنائية وهو ما يسمح فقط بستة عشر رمزًا مختلفًا (اثنين مرفوعة لأس أربعة) وفي وقت تالٍ وصل حيز الكلمة الرقمية إلى ستة أرقام ثنائية مما يسمح فقط بأربعة وستين رمزًا (اثنين مرفوعة لأس ستة)، وهذا بالكاد يستوعب الرموز المطلوبة للغة واحدة (كالإنجليزية مثلاً) إضافة إلى رموز تحكم ضرورية أخرى، مما يعني صعوبة استيعاب أكثر من لغة معًا في آن واحد.

٢. ارتباط البرمجيات (السُّوفْت-وِير) عضويًّا واعتمادها مباشرةً على عتاد (هارْد-وِير) تلك الحاسبات وتشكيلهما معًا منتجًا واحدًا تقوم بتصنيعه نفس الشركة، وبالتالي فإن تغيير لغة تفاعل تلك البرمجيات كان يحتاج إلى تعديل المنتج وخط

إنتاجه. فإذا كان للشركة المنتجة وقتها أن تدعم عدد س من اللغات فإنه كان عليها أن تقيم وتدير س من خطوط الإنتاج لما تبنيه من حواسب بعتادها وبرمجياتها وهو أمر ليس باليسير.

٣. انغلاق أنظمة هذه الحاسبات وبِنَاها أمام المطورين من خارج الشركات المصنعة لها، مما جعل من العسير عليهم أن يعدلوا برمجيات تلك الحاسبات.

مع بداية النصف الثاني من ثمانينات القرن العشرين الميلادي وقعت تطورات هامة في عالم الحاسبات الصغيرة حيث صارت المشغلات الدقيقة بكلمات يبلغ حيزها ثمانية أرقام ثنائية هي السائدة مقابل أسعار معتدلة[١]. وبدأ خطان متوازيان في صناعة الحاسبات الصغيرة؛ فنشأت شركة «آبـِّل» (Apple) الأمريكية وصنعت حاسبات «الماكِينْتُوش»[٢] الصغيرة بنظام تشغيل متماسك «ماك-أو-إسّ» يفصل الارتباط العضوي بين العتاد وبين البرمجيات، كما طورت الشركة برمجيات تميزت في عالم الرسوميات «الجرافِيكْس» وتنسيق المطبوعات وإعدادها للنشر والطباعة فيما عُرِفَ وقتها باسم «النشر المكتبي»، ونجحت في نقل عالم نشر الصحف والمجلات من عصر الميكنة إلى عصر الحوسبة، وأمكن الاستجابة للطلب على إتاحة هذا النشر المكتبي بلغات العالم الكبرى المختلفة، وفي أقل من سنتين كان النشر المكتبي بالخط العربي متاحًا بصورة جيدة على هذا النوع من الحاسبات مدعومًا بسوق الصحافة الخليجية السخي. أما على الخط الموازي الآخر ولكن على نطاق انتشار أوسع كثيرًا طرحت «آي-بي-إم» العملاق الأمريكي في عالم صناعة الحاسبات حاسبَها الصغيرَ تحت اسم «الحاسب الشخصي» (Personal Computer- PC) بسعر جذاب[٣] وألقت وراءه ثقلها إدراكًا منها أن مستقبل سوق الحاسبات سوف تسيطر عليه الحاسبات الصغيرة وليست العملاقة، وحقق هذا النوع من الحاسبات الذي كان يعمل على مشغلات دقيقة من شركة «إنْتِل» بكلمات يبلغ

١- كانت هناك أيضًا في ذلك الوقت مشغلات دقيقة بكلمات يبلغ حيزها ستة عشر رقمًا ثنائيًّا لكنها باهظة الثمن كانت تبنى عليها حاسبات أرفع أداءً تستخدم في تطبيقات متخصصة لكنها باهظة الثمن كانت يسمى كل منها "محطة عمل Work Station".

٢- الماكينتوش نوع من التفاح المنتشر في أمريكا الشمالية.

٣- كان يمكن شراؤه بأقل من ألف وخمسمئة دولار أمريكي في حدود عام ١٩٨٧م انخفضت لما دون الألف دولار أمريكي بنهاية ذلك العِقد.

حيزها ثمانية أرقام ثنائية نجاحًا هائلاً وبيعت منه ملايين الوحدات، وكان من أسباب نجاحه تزويده بنظام تشغيل «دُوسْ» (DOS) من شركة «مايِكْروسوفت» حيث كان نظامًا مفتوحًا كرس مبدأ الفصل بين تطوير البرمجيات وبين تصنيع العتاد في مثل ذلك النوع من الحاسبات الصغيرة فسمح بذلك للمطورين من خارج كلتا الشركتين بتطوير برمجيات لهذه الحاسبات. وضمن هذا المناخ شرعت شركات عربية في محاولات إنتاج برمجيات عربية لتلك الحاسبات وشرعت أخرى في محاولة تعريب بعض البرمجيات الشهيرة ولاقت تلك المحاولات نجاحًا أثبت أن التعريب أمرٌ ممكن.

وبعد عام ١٩٩٠م حين طرحت «مايِكْروسُوفْت» نظام تشغيلها «ويندوز – الإصدار ٣,١» بواجهة تشغيل رسومية ارتفعت موجة انتشار الحاسبات الشخصية لتشمل كل نواحي العالم تقريبًا ولتحتكر «مايكروسوفت» سوق أنظمة التشغيل لهذا النوع من الحاسبات الذي كانت إمكاناته الحاسوبية بحلول ذلك الوقت قد ارتقت كثيرًا مما دفع «مايِكْروسُوفْت» لطرح ملحقات لنظام تشغيلها تدعم لغات العالم المختلفة ومن بينها العربية سواء في واجهة استخدام نظام التشغيل أو في برامجها المكتبية – خاصة برنامجها الأشهر لكتابة الوثائق وتنسيقها «وُورْد» – عرضًا على الشاشة وطباعةً على الورق. وبطبيعة الحال لم يخل الأمر في البداية من بعض المشكلات والمنغصات الهامشية التي أخذت تتلاشى تدريجيًّا مع الإصدارات التالية المتلاحقة لنظام التشغيل «ويندوز».

والآن، وبعد شيوع معيار «النِّظام المَوَحَّد لشفرات الحُرُوف» (Unicode) ذي الستة عشر رقمًا ثنائيًّا لتمثيل النصوص والذي يتسع لحروف معظم لغات العالم المعاصرة واسعة التداول (ومن بينها العربية بالطبع)، فإن الحرف العربي مدعوم بسلاسة وبشكل تلقائي – دون الحاجة لتنصيب ملحقات – على كافة أنواع الحاسبات تقريبًا سواءً أكانت الأجيال المتقدمة من الحاسبات الشخصية وصولاً إلى الحواسب الكفية المدمجة مع الأجيال الحديثة من الهواتف النقالة، وغير ذلك من الأصناف المتعددة للحاسبات.

وتجب الإشارة في ختام هذا الاستعراض التاريخي إلى أن ميكنة وحوسبة الخِطاطة العربية كانت كلتاهما وما زالتا تتعاملان في مسارهما الرئيسي أساسًا مع الخطوط العربية المنتظمة، ونعني بها تلك الخطوط التي تنساب كتابتها في اتجاه واحد؛ وهو الاتجاه أفقيًّا من اليمين إلى اليسار، وفي نفس الوقت تسمح برسم أي نص عربي عبر انتقاء سلسلة

من الوحدات التي تنتمي إلى فئة مغلقة من الوحدات الرسومية «الجرافيمات». وفي حين أن هذا التعريف يستبعد الخطوط الزخرفية، فإن إضافة الوضوح والمعيارية إليه يجعل خطوط النسخ البسيطة (راجع الشكل رقم ١-٥) هي محل التركيز الأكبر لهذه الصناعة.

٣- تحديات الخِطاطة العربية التي تواجه التعرف عليها آليًّا

بعد الاستعراض التاريخي لنشأة الخِطاطة العربية ثم ميكنتها وحوسبتها، نقوم فيما يلي بالتدقيق في بعض تفاصيل هذه الخِطاطة التي تشكل تحديات ينبغي على أي نظام ناجح للتعرف الآلي على النص العربي المكتوب أن يعالجها بفعالية.

١. اتصال الحروف: يمكن كتابة اللغات الأوروبية بحروف منفصلة (وهي السائدة في النسخ والطباعة لبساطتها) أو بحروف متصلة (وتستخدم كثيرًا عند الكتابة بخط اليد)، في حين أن لغات أمم كبرى في أقصى شرق آسيا (كالصينية واليابانية والكورية) تُخَطُّ دائمًا بالحرف المنفصل، أما الخِطاطة العربية (ويشاركها في ذلك البنغالية والهندية) فتُكْتَب دائمًا بالحرف المتصل.

ومن وجهة نظر أية آلية حاسوبية للتعرف على الأنماط الرسومية، فإنه مع تثبيت جميع الظروف الأخرى يكون التعرف على الأنماط وهي منفصلة أيسر من التعرف عليها وهي متصلة ببعضها البعض، حيث يتوجب في حالة الاتصال حل مسألة تعيين حدود كل رمز رسومي «جرافيم» (وهو ما يطلق عليه الباحثون في هذا المجال اسم «التقطيع» (segmentation)) إضافةً بالطبع إلى مسألة «التعرف» (recognition) على الحرف الذي يرمز إليه كل جرافيم. وكما يمكن استشفافه من الشكل رقم (١-٧) أدناه، فإن التعرف الصحيح على الرموز يتطلب تعيين حدودها بشكل سليم، ولكن تعيين هذه الحدود بدوره يتطلب معرفة الرموز أولاً!

وهكذا، فإنه لا مفر من حل كلتا المسألتين («التعرف» و «التقطيع») آنيًّا وهو ما يضاعف التحدي، ويتناول أول أقسام المبحث الثالث من هذا الفصل المزيد عن أساليب تحقيق ذلك. بقي أن نؤكد ثانيةً على أن الخطوط العربية المنتظمة حسب التعريف الذي أوردناه في نهاية القسم السابق هي تلك التي يمكن تعيين

حدود وفواصل واضحة بين جرافيماتها بينما لا يمكن عمل ذلك مع الخطوط غير المنتظمة؛ مثل تلك الزخرفية (راجع شكل رقم ١-٤ أعلاه).

الترجمة وسيلة أساسية لتبادل الحضارات بين الشعوب على مر العصور

الشَّكل ١-٧: الاتصال الأفقي بين الحروف المتتالية في خطوط النسخ وتعيين حدوده.

٢. التداخل بين حدود الجرافميات: مما يزيد من التحدي السالف عرضه في النقطة السابقة أننا نرى أحيانًا بعض التداخل (أو التراكب) الطفيف بين حدود جرافيمات بعض الخطوط العربية المنتظمة كما يمثل له الشكل رقم (١-٨) أدناه.



الشَّكل ١-٨: مثال على التراكب بين الحروف المتتالية.

٣. تغير رسم الحرف مع تغير موضعه في الكلمة: وهذا التغير هو بالطبع نتيجة للكتابة متصلة الحروف، ويؤدي هذا التغير إلى زيادة كبيرة عدد الرموز الرسومية التي يتوجب أن يتعامل معها أي نظام للتعرف الآلي على النص العربي المكتوب بالمقارنة مثلاً مع اللغات الأوروبية التي تُخَط عند طباعتها عادة بالحرف المنفصل.

ومن وجهة نظر أية آلية حاسوبية للتعرف على الأنماط الرسومية، فإنه مع تثبيت جميع الظروف الأخرى يكون التعرف أكثر دقة كلما قلت عدد الأنماط الرسومية المختلفة والعكس بالعكس.

على ، العربية ، مع ، قطاع

الشَّكل ١-٩: مثال على اختلاف رسم الحرف العربي مع اختلاف موقعه من الكلمة.

٤. الجرافيمات المركبة من أكثر من حرف واحد: الكثير من خطوط النسخ المستخدمة بكثافة في ميكنة وحوسبة نسخ وطباعة الكتابة العربية (ومن بينها الخَطُّ الذي كُتِبَ به هذا الكتاب) تحتوي على العديد من الجرافيمات المركبة من حرفين أو ثلاثة (والتي لا يمكن التعامل معها إلا كوحدة رسومية واحدة)، ويبين شكل رقم (١-١٠) أدناه أمثلة على بعض هذه الجرافميات المركبة.

وعلاوةً على ما أشرنا إليه في النقطة السابقة، فإن هذه الجرافيمات المركبة ترفع عدد الأنماط الرسومية التي يتوجب على أي نظام للتعرف على النص المكتوب التعامل معها مما يرفع درجة الصعوبة، وإذا كانت الخِطاطة الإنجليزية – على سبيل المثال – في هذا الصدد قد تكتفي بحوالي ثمانين جرافيمًا فإن العربية تحتاج إلى ما يزيد على مئة وتسعين.

مـــ ، لجـــ ، لا ، لمحـــ

الشَّكل ١-١٠: أمثلة على جرافيمات مركبة من أكثر من حرف في أحد خطوط النسخ.

٥. النَّقْط: كلما كانت الفروق الشكلية بين الأنماط الرسومية أكبر، فإنه مع تثبيت جميع الظروف الأخرى ترتفع قدرة أي نظام للتعرف الآلي عليها. وبما أن نسبة كبيرة من جرافيمات الخطاطة العربية متشابهة شكليًّا إلى حد بعيد ولا تتمايز إلا بوجود أو غياب النقاط – أو الهمزة في بعض الأحيان – فإن ذلك بالتأكيد يرفع التحدي أمام أي نظام للتعرف الآلي على النص العربي المكتوب.

بـــ ، نـــ ، تـــ ، يـــ ، ثـــ | ح ، خ ، ج | س ، ش | ط ، ظ

الشَّكل ١-١١: أمثلة على مجموعات الحروف العربية التي تتمايز بالنقاط.

٦. علامات الضبط الصوتي (التشكيل): تعتبر علامات التشكيل في الخِطاطة العربية تعقيدًا إضافيًّا أمام أي نظام للتعرف الآلي على النص العربي المكتوب وذلك لأنها لا تقع في سياق تسلسل أفقي مثل الجرافيمات الهجائية ولكنها في مواضع رأسية فوقها أو تحتها (كما يبين شكل رقم ١-١٢ أدناه).

ولذلك فإن مثل هذه الأنظمة تتعامل مع علامات التشكيل إما بمحاولة اكتشافها مبكرًا ثم حذفها قبل عملية التعرف على النص المكتوب، أو بتجاهلها على اعتبار أن الكتابة العربية المعاصرة نادرًا ما تضاف إليها هذه العلامات إلا لأغراض تعليمية أو عند اقتباس النصوص الدينية أو التراثية.

اَلتَّرْجَمَةُ وَسِيلَةٌ أَسَاسِيَّةٌ لِتَبَادُلِ الْحَضَارَاتِ بَيْنَ الشُّعُوبِ

الشَّكل ١-١٢: مثال على عبارة عربية بعلامات الضبط الصوتي الكامل.

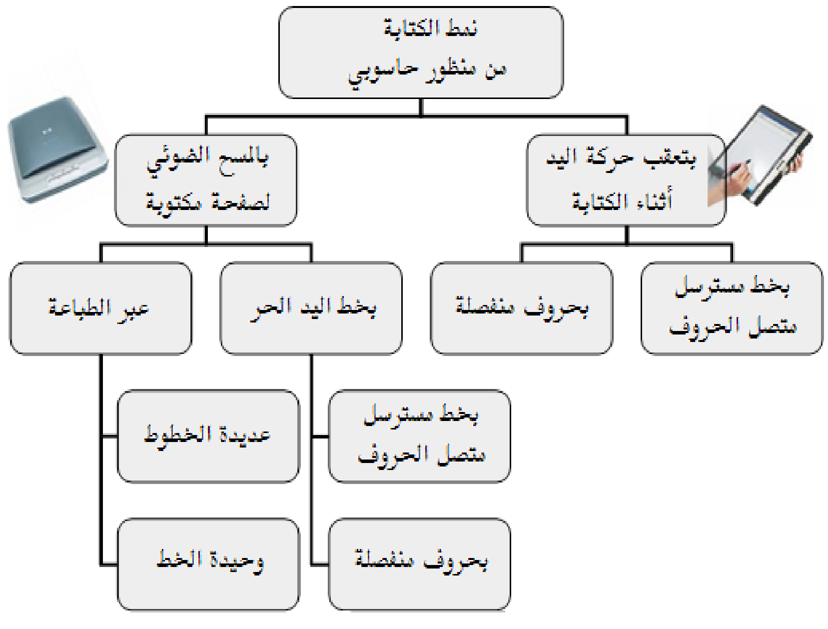
المبحث الثَّاني
أنواع التَّعرُّف الآليّ
على النَّصّ العَرَبيّ المكتوب

١- تصنيف أنماط الكتابة وتمثيلها من منظور حاسوبي.

٢- تطبيقات التعرف الآلي على النص العربي المكتوب.

١- تصنيف أنماط الكتابة وتمثيلها من منظور حاسوبي

تجري «رَقْمَنَةُ[١]» (Digitization) الكتابةِ ومن ثَمَّ تمثيلُها حاسوبيًّا حَسْبَ ظروف إنتاج الكتابة المستهدَفة وذلك وَفْقَ المخطَّط الذي يلخصه الشكل رقم (١-١٣) أدناه:

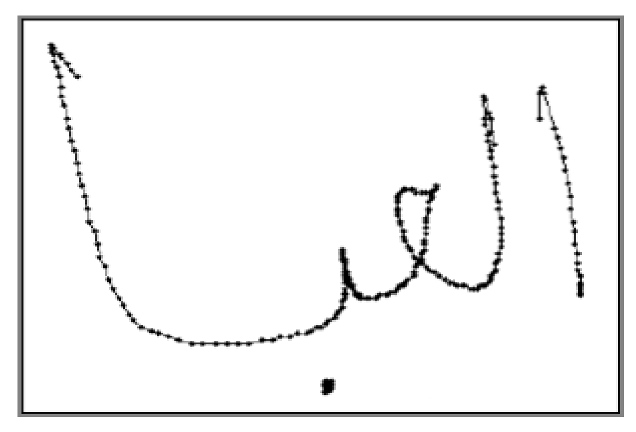


الشَّكل ١-١٣: تصنيف أنماط إنتاج الكتابة حسب التعامل معها حاسوبيًّا.

فهناك الكتابة اليدوية حيث يجري تعقبُ حركة اليد أثناءَ مرور الأصبع أو قلم خاص على لوح حساس لِلَّمْس سواء أكان لوحَ كتابةٍ أم شاشةً لحاسبٍ كَفِّيٍّ أم هاتفًا ذكيًّا أم شاشةً لحاسبٍ لَوْحيٍّ ... إلخ، ويتم هذا التعقب عبر تعيين والتقاط الموضع النسبي لليد على اللوح الحساس لِلَّمْس كلَّ لحظة[٢] وإرسالِها إلى ذاكرة الحاسب ويشكل تتابعُ هذه اللقطات التمثيلَ الرَّقْمِيَّ للكتابة اليدوية كما يبينه الشكل رقم (١-١٤) أدناه، وهو ما يسمح بتكوين «المنحنيات المناظرة لمسار الكتابة» (Writing Contours) وهي المعلومة الأساسية في عملية التعرف على هذه الكتابة.

١- أي تحويلها من حالتها الأصلية «التناظرية Analog» إلى الحالة «الرقمية Digital» من أجل إدخالها ومعالجتها بواسطة الحاسبات الرقمية.

٢- تستطيع التقنيةُ الحاليةُ رصدَ عدةِ مئاتٍ من اللقطات لموضع اللَّمْس على اللوح الحساس في كل ثانية، وهو ما يضمن تمثيلاً سَلِسًا للكتابة لا تُفْقَدُ فيه تفاصيلُ حركة اليد.



الشَّكل ١-١٤: مثال على رقمنة الكتابة اليدوية أثناء كتابة كلمة «العبارة»
وتمثل العقد الصغيرة مواضع اليد عند اللحظات التي جرى تسجيلُها.

وقد تكون الكتابة اليدوية تدوينًا مسترسلاً بحروف متصلة كما في الكتابة العادية الحرة (كما في شكل رقم ١-١٤ عالِيَه) وهي الحالة الأصعب في التعرف عليها، أو تكون مجرد حروف أو علامات منفصلة وهي الحالة الأيسر في التعرف عليها، ونتناولُ بَعْضَ التطبيقات المتصلة بكلتا الحالتين في القسم التالي.

وعلى الناحية الأخرى هناك الصفحات المحتوية على نصوص مكتوبة بالفعل، حيث يجري تحويل كُلٍّ من هذه الصفحات إلى «صورة رقمية» (Digital Image) عبر جهاز «الماسح الضوئي» (Scanner)، وفي هذا الصدد يُؤخَذ الآن بالمعيار الأوسع قَبولاً لعملية المسح الضوئي بتباين أحادي اللون (بالأبيض والأسود) وبتحليل طولي يساوي ستمئة نقطة في كل بُوصَة[1] طولية (أي ثلاثمئة وستين ألف نقطة في البُوصَة المربَّعة) [2].

وقد تكون الصفحة الممسوحة ضوئيًّا مكتوبةً بخط اليد الحر سواء أكانت كتابة مسترسلة بحروف متصلة (كما يبين الشكل رقم ١-١٥ أدناه مثالاً على ذلك) وهي الحالة الأصعب في التعرف عليها، أو تكون مجرد حروف أو علامات منفصلة وهي الأيسر في التعرف عليها، ونأتي على ذِكْر بعض التطبيقات المتصلة بكلتا الحالتين في القسم التالي.

١- البوصة تساوي ٢,٥٤ سنتيمترًا تقريبًا.

٢- كان المعيار السابق هو ثلاثمئة نقطة في كل بوصة طولية (أي تسعين ألف نقطة في البوصة المربَّعة).

الشَّكل ١-١٥: مثال على الصورة الرقمية الناتجة عن المسح الضوئي لصفحة تحتوي على كتابة بخط اليد، وتمثل محتويات كل مربع صغير إحدى نقاط هذه الصورة الرقمية

وعلى الناحية الأخرى قد يكون النص في الصفحة الممسوحة ضوئيا مطبوعا (كما يبين الشكل ١-١٦ أدناه مثالاً على ذلك) ويتصاعد تحدي التعرف على النص عندما تتعدد الخطوط (الأنماط) المستخدمة في طباعته عما إذا كان مطبوعًا بخط واحد فقط.

الشَّكل ١-١٦: مثال على الصورة الرقمية الناتجة عن المسح الضوئي لصفحة تحتوي على نص مطبوع، وتمثل محتويات كل مربع صغير إحدى نقاط هذه الصورة الرقمية

وبصفة عامة فإن التعرف الآلي على النصوص المخطوطة يدويًّا أصعب كثيرًا من التعرف على النصوص المطبوعة، ونتناولُ بَعْضَ التطبيقات المتصلة بكلتا الحالتين في القسم التالي.

٢- تطبيقات التعرف الآلي على النص العربي المكتوب

تكتسبُ تقنيةُ التعرف الآلي على النص عبر تعقُّب حركة اليد أثناء الكتابة أهميةً صناعيةً تجاريةً متعاظمةً تتناسب مع الانتشار الهائل للأجيال الجديدة من الهواتف الذكية التي يحتوي معظمُها على شاشات حساسة لِلَّمْس وكذلك انتشار «الحاسبات اللَّوْحية» (Tablet Computers) وجميعها ذات شاشات حساسة لِلَّمْس، وذلك حيث أن هناك مَيْلاً إلى استبعاد لوحة المفاتيح من تصميم هذه الأجهزة واستخدام اللمس بالأصابع أو بقلم مخصوص للتفاعل معها وإدخال البيانات إليها. ويتفاوت التحدي في التعرف على هذه البيانات بدءًا من لمس أحد الخيارات المتعددة، مرورًا بكتابة مجرد علامة أو حرف أو رقم منفصل، وصعودًا إلى الكتابة الحرة المسترسلة بحروف متصلة. ويلاحَظ أن برمجيات التعرف على هذه الأنواع من البيانات غالبًا

ما تحتوي على أدواتٍ مساعِدةٍ للمستخدم تمكنه من التحقق من صحة التعرف على ما يكتبه والتدخل لتصويبه عند اللزوم، بل تحتوي نُوَاةُ هذه البرمجيات أحيانًا آلياتِ تدقيقٍ إملائي لتوجيه عملية التعرف أو لتصويب ما جرى التعرفُ عليه من كتابة عربية مسترسلة بحروف متصلة، كما أنها تتضمن أدواتٍ «لِلتَّكيُّف» (Adaptation) مع خط المستخدم، وتتعاضدُ كلُّ أو بعضُ هذه الأدوات للارتقاء بأداء التعرف على هذا النوع من الكتابة إلى مستوًى يُرْضِي المستخدم. [٤، ٧، ٨]

وتجدرُ - إضافةً إلى ما سبق - الإشارةُ إلى التطبيقات التعليمية والتربوية لهذه الفئة من برمجيات التعرف الآلي على النصوص حيث تُستخدَم في المساعدة على تعلُّم كتابة اللغة للنشء أو للكبار أو على تحسين الخط؛ حيث تعرض مثلُ هذه البرمجيات مسارات الكتابة اليدوية النموذجية لأمثلة من النصوص وتطلب من المستخدم أن يتتبع هذه المسارات بقلم مخصوص في يده على لوح حساس لِلَّمْس أثناءَ مرحلة التدرُّب، كما يُطْلَب أثناءَ مرحلة تقويم الأداء من المستخدم أن يخط على اللوح الحساس أمثلةً نصيةً، ومن ثَمَّ تقومُ هذه البرمجياتُ بمقارَنة مسارات الكتابة اليدوية للمستخدم مع المسارات النموذجية لكتابة أمثلة التدريب أو تقويم الأداء وبِناءً على نتيجة المقارَنة يُفادُ المتعلم بالتصحيحات المطلوبة لتحسين أدائه [١].

أما على جانب تقنيات التعرف على الكتابة المسترسلة في الصفحات الممسوحة ضوئيًّا سواء أكانت مخطوطةً باليد أم مطبوعة، فإن التطبيق الأساسي لها هو العمل على الرقمنة الآلية لتراث البشرية الضخم من المواد النصية الموجودة بالفعل على الأوراق[١] مقابلَ تكلفةٍ مالية وزمنية وإدارية أقل كثيرًا مما يستغرقه إنجازُ تلك الرقمنة بأيادي جيوش من «الضاربين على لوحات المفاتيح» (Typists). وتوفر رقمنة هذا التراث النصي مزايا كبرى تتمثل في:

١- من أمثلة المشروعات الكبرى في هذا الصدد «مشروع رقمنة المليون كتاب» الذي تقوم عليه شركة «جوجِل»؛ انظر المرجع رقم [١١] من مراجع هذا المبحث.

١. إمكانية المحافظة على الوثائق إلكترونيًّا بتكلفة ضئيلة جدًّا بالمقارَنة مع تكلفة حفظها ورقيًّا[١]، ولمدة أطول بكثير من عمر هذه الوثائق في صورتها الورقية يمكن أن تصل إلى ما شاء الله؛ حيث يمكن نقل الوثيقة من وسيطِ تخزينٍ إلكترونيٍّ قديمٍ إلى آخر جديد دون فقدان أي شيء من محتوياتها قبل انتهاء العمر الافتراضي للوسيط القديم، وباحتمالات ضياع متضائلة إلى ما شاء؛ حيث يمكن استنساخ الوثيقة إلكترونيًّا عددًا غير محدود من المرات على وسائط تخزين متعددة تُحْفَظ في أماكن مختلفة[٢].

٢. يؤدي تحويلُ الوثائق إلى صور رقمية عبر مسحها ضوئيًّا ثم تحويلُها بعد ذلك إلى نصوص رقمية عبر تقنيات التعرف الآلي على النصوص إلى إمكانية تشغيل تقنيات معالجة النصوص على محتويات هذه الوثائق وخصوصًا تلك التي تستخلص المعلومات وتستنبط المعرفة من الأوعية النصية؛ ومن أهمها «محرِّكات البحث النصي» (Text Search Engine) و«استرجاع البيانات» (Information Retrieval- IR) فضلاً عن «التنقيب المعلوماتي في النصوص» (Text Mining)[٣] وكذلك «تصنيف الوثائق» (Document Classification) و«تلخيصها» (Document) Summarization)، وقد

١- بأسعار نهاية عام ٢٠١١م يتكلف القرص الصلب الخارجي سعة تخزين واحد تِيرًا (مليون مليون) بايتْ حوالي خمسين دولارًا أمريكيًّا وبافتراض أن الكتاب متوسط الحجم يحتاج سعة مئة مِيجَا (مئة مليون) بايت لتخزينه بما فيه من صور بدرجة عالية الجودة وبعمر افتراضي لا يقل عن خمس سنوات تحت ظروف قاسية، فإن هذا القرص يسع لتخزين لعشرة آلاف كتاب وهذا يعني أن الكتاب الواحد يتكلف نصف سِنْت أمريكي. وعلى الجانب الآخر فإن المتر المربع السطحي في غرفة مكتبة مجهزة على مستوى جيد يتكلف في نفس التاريخ ما لا يقل عن مئتين وخمسين دولارًا في العام ويسع ما لا يزيد عن ألف كتاب من نفس الحجم وهو ما يعني أن الكتاب الورقي يتكلف تخزينه لمدة خمس سنوات في حالة جيدة حوالي مئة وخمسة وعشرين سنتًا أمريكيًّا أي مئتين وخمسين ضعف التكلفة الإلكترونية. هذا مع ملاحظة أن تكلفة الحفاظ على المخطوطات التاريخية تتجاوز ذلك بكثير نظرًا لاحتياجها إلى تجهيزات خاصة تحافظ على مادتها من البِلَى مع مرور الزمن.

٢- من الإنصاف أن ننتبهَ إلى قضية التحقق من «أصالة Authenticity» الوثائق الرقمية حيث لا توجد حتى الآن - رغم الأبحاث الجارية على هذه المسألة - وسيلةٌ حاسمة للتأكد من عدم التلاعب إلكترونيًّا بمحتويات الوثيقة المرقمَنة سوى المطابَقة بجسم الوثيقة الأصلية!

٣- تهدف آليات التنقيب المعلوماتي بصفة عامة في أي وعاء للبيانات (سواء أكانت نصوصًا أم سواها) إلى اكتشاف أنماط التكرار والارتباط السائدة بين مفردات هذه البيانات مما ينبئ عن احتمال قوي لوجود معلومة يعبر عنها هذا النمط أو ذاك، وتُعتَبَرُ لذلك أداةً قويةً لتحويل البيانات (كمادة خام قليلة الفائدة) إلى معلومات (عالية القيمة).

تُجْمَع كُلُّ هذه التقنيات مع أدوات أخرى تقليدية لإدارة البيانات والمعلومات ضمن أنظمة متكاملة يُصطَلَح على تسميتها في المجْمَل «نُظُم إدارة الوثائق» (Document Management Systems- DMS).

٣. ومن الأهمية بمكانٍ فيما يتعلق برقمنة نصوص الوثائق أن نضع الأداء البشري على محك المقارَنة بمستويات الأداء الراهن لتقنيات التعرف الآلي على النصوص العربية المكتوبة من الوثائق المسموحة ضوئيًّا. فالضارب المحترف على لوحة المفاتيح يستطيع في المتوسط نَسْخَ حوالي ستين كلمة عربية في الدقيقة الواحدة بمعدل خطأ للكلمات (Word Error Rate: WER) لا يتجاوز الثلاثة في المئة كما يُمْكِنُ من خلال مراجعةٍ يدويةٍ النزولُ بمعدل الخطأ للكلمات إلى ما دونَ النصف في المئة[١]، ومع متوسط أجور لهذا النوع من العمالة - بمعدلات نهاية عام ٢٠١١م - يبلغ نحو دُولارَيْن أمريكيين في الساعة فإن التكلفة الزمنية لرقمنة ألف كلمة بجودة عالية تكون في حدود نصف ساعة عمل كما تكون التكلفة المالية لذلك في حدود دولارَيْن أمريكيين[٢].

أما تقنيات التعرف الآلي على النصوص؛ فإن التكلفة الزمنية ليست عائقًا حيث يتكفل بها كُلٌّ من التسارعِ المطَّرِد لقدرات الأجيال المتعاقبة من «المعالجِات الدقيقة» (Microprocessors)[٣] التي تُشغِّل الحاسبات الرقمية، وكذلك إمكانيةِ رَفْع سرعة التعرف على النصوص بالقدر المطلوب عبر أيٍّ من البِنَى المتعددة للحوسبة «الموزَّعة» (Distributed Processing) أو «المتوازِية» (Parallel Processing)[٤]. كما أن التكلفة المالية كذلك ليست عائقًا حيث أنه مهما ارتفعت تكاليف البحث والتطوير والتشغيل

١- هذا بافتراض خُلُوِّ الوثيقة الجاري نسخُها - سواء أكانت مطبوعة أم مخطوطة باليد - من التشوهات أو الشوشرة البصرية العالية التي تؤدي إلى الالتباس في قراءة المحتويات النصية لها.

٢- هذا بافتراض العمل على كميات ضخمة من النصوص ضمن خط إنتاج تتوازى فيه عمليتا النسخ والمراجعة.

٣- وذلك وفْقَ ما يُعْرَف «بقانون مُورْ Moore's Law» وهو عبارةٌ عن استقراء عملي لواقع تطور الحواسيب مع الزمن وينص على تضاعف القدرات الحاسوبية - من معالجَة وتخزين ... إلخ - مرةً كُلَّ ثمانية عشر شهرًا تقريبًا.

٤- وأيًّا كانت البِنْية الموزَّعة أو المتوازية فإن المبدأ البسيط وراءَها هو تشغيل حواسيب عديدة بدلاً من حاسوب واحد فقط من أجل الانتهاء من نفس المسألة في زمن أقل..

لأي نظام ناجح للتعرف الآلي على النصوص، فإن تكلفة التعرف على كل ألف كلمة تتضاءل باطراد في تناسب عكسي مع إجمالي النصوص التي جرى التشغيلُ عليها. ويبقى العائقُ الحقيقي متمثِّلاً في دقة التعرف لهذه النُّظُم؛ حيث يعرض الجـدولُ رقم (١-١) أدناه هوامشَ الخطأ للكلماتِ المِقيسةَ لأفضل ما جرى تطويرُه من أنظمة التعرف الآلي على النصوص العربية [٣، ٩، ١٠] في ظل مستوياتٍ متدرِّجةٍ من جودة الصورة الممسوحة ضوئيًّا وكذلك في ظل مستوياتٍ متدرجةٍ من التنوع في الوحدات الرسومية للكتابة (أي «الجِْرافِيمات»)؛ ومن الواضح أن هامش الخطأ يتسع باطراد مع تصاعد مستوى «الشوشرة» (Noise) البصرية التي تشوب الصور كما يتسع هذا الهامش مع تزايد التنوع والتباين في جرافيمات الكتابة.

٤. وتفيد الدراسات التي حللت أداء تقنيات استخلاص المعلومات واستنباط المعرفة - المشار إليها في الفقرة ب عالِيَه - عند تشغيلها على أوعية بها «نصوص متدهورة» (Degraded Text) (أي نصوص مَعِيبة ذات نسبة معتبَرة من الأخطاء) أن مردود هذه التقنيات يظل فعالاً إذا لم تتجاوز نسبة التدهور الخمسة عشر في المئة [٥] (وترفع بعض الدراسات هذه النسبةَ إلى عشرين في المئة[1] مع ملاحَظة أنه بعد إجراء عملية استخلاص المعلومات أو استنباط المعرفة التي يطلبها المستخدمُ يجري عادةً عَرْضُ الوثائق الواردة في نتائج العملية المطلوبة على هيئة الصورة الضوئية الممسوحة -وتُسَمَّى نسخةُ الاطِّلاع - وليس على هيئة النص الرقمي الذي جرى التعرفُ عليه آليًّا لاحتوائه على نسب الأخطاء في الكلمات المشار إليها.

وعلى ذلك فإن الحالات المظللة بالرمادي في الجدول هي فقط التي تصلح لتطبيق تقنيات استخلاص المعلومات واستنباط المعرفة عليها بفعالية مقبولة.

١- في حين يزعم بعضُ الباحثين أن أداء مثل تلك التقنيات قد لا ينهار بالكامل حتى مستوياتِ خطأٍ عشوائية للكلمات تبلغ الخمسين في المئة!

التنوع - الجودة ↓ →	طباعة ليزر أو وثيقة أصلية ✩✩✩✩	صفحات كُتُب ✩✩✩	نسخ ضوئي لمرة واحدة ✩✩	نسخ ضوئي مرتين أو صفحات جرائد ✩
مطبوعة بخط واحد وحجم واحد ✩	٢,٥٪	٣٪	٨٪	حوالي ١٥٪
مطبوعة بمختلف الخطوط والأحجام ✩✩✩	١٣٪	١٤٪	٢٥٪	أكبر من ٤٥٪
مخطوطة باليد «بعناية ووضوح» ✩✩✩✩✩	بعد التدرب على خط الكاتب ٢٥٪	×	أكبر من ٥٠٪	أكبر من ٧٥٪

الجدول ١-١: هوامش الخطأ المقيسة في أداء أفضل ما جرى تطويره - حتى نهاية ٢٠١١م - من أنظمة التعرف الآلي على النصوص المكتوبة في صفحات ممسوحة ضوئيًا، وذلك في ظل ظروف متدرجة من جودة الصور الممسوحة ضوئيًّا وكذلك التنوع في جرافِيات الكتابة.

٥. وبينما لا تزال مخرَجاتُ التعرف الآلي على النصوص العربية المخطوطة باليد في كل الحالات خارجَ إطار الاستفادة من تقنيات استخلاص المعلومات واستنباط المعرفة نظرًا للاتساع الكبير في هامش الخطأ بها، فإن تركيز الأبحاث التطبيقية في هذا الصدد يتركز على إنجاز تطبيقات ضيقة النطاق؛ إما من جهة التقيد بحصيلة لغوية محدودة جدًّا تحصر البحثَ داخل نُظُم التعرف وتكفي لمهام مثل التعرف على عناوين البريد أو قوائم مختصرة بأصناف سلع ... إلخ، أو من جهة التقيد بالرموز والأرقام والعلامات والحروف المنفصلة في مهام مثل تصحيح الاختبارات الموضوعية أو قيمة الصكوك النقدية (أي «الشيكات Checks /Cheques») ... إلخ.

المبحث الثَّالث
أساليب التَّعرُّف الآليّ على النَّصّ العربيّ المكتوب

١ – التعرف على النص المكتوب كأحد أنظمة التعرف على الأنماط.

٢ – أساليب التعرف على الكتابة العربية بتعقب خط اليد.

٣– أساليب التعرف على الكتابة العربية المطبوعة.

٤ – أساليب التعرف على الكتابة العربية المخطوطة يدويًّا.

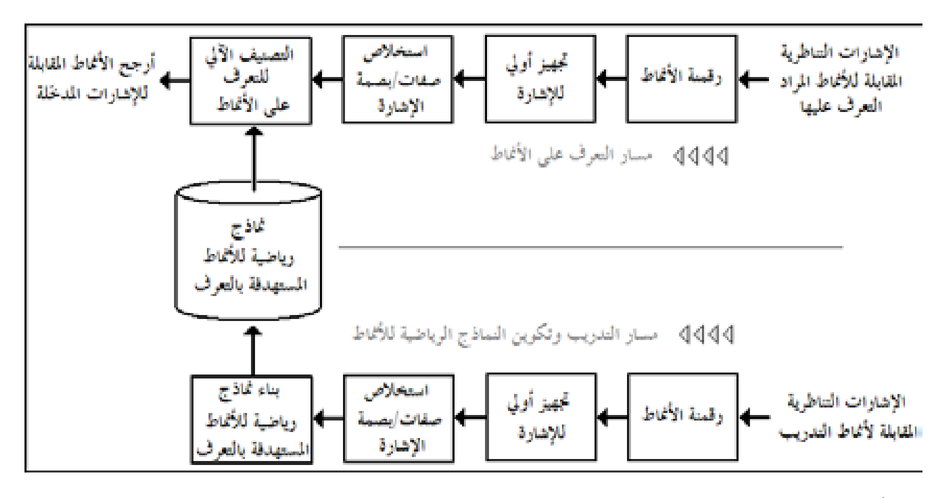
٥ – بناء الموارد اللغوية لتدريب وتقويم أنظمة التعرف على الكتابة العربية.

١ - التعرف على النص المكتوب كأحد أنظمة التعرف على الأنماط

تنتمي أنظمة التعرف الآلي على النص المكتوب إلى مجال أرحب من علوم الحاسبات التطبيقية ألا وهو «التعرف الآلي على الأنماط» (Pattern Recognition)، ويمكن وضعُ الهيكل الوظيفي لهذه النظم في الإطار العام المبين في الشكل رقم (١-١٧) أدناه [٦]؛ حيث يتم تحويلُ «الإشارات التناظرية» (Analog Signals) المقابلة للأنماط المطلوب التعرف عليها (وهي النصوص المكتوبة/ المخطوطة في حالتنا هذه) إلى «إشارات رقمية» (Digital Signals) - كما ورد خلالَ القسم الأول من المبحث السابق - ومن ثَمَّ تزويدُ الحاسوب بها، ثم يجري «تجهيزٌ ابتدائي» (Preprocessing) لهذه الإشارات الرقمية (وذلك مثلُ استبعاد بعض الأنساق النمطية من الشوشرة)، ثم يتم بعد ذلك استخلاصُ بصمات هذه الإشارات (أي مجموعة الصفات الرياضية المميِّزة لها).

ويمتاز بعد ذلك مسار التدريب حيث يجري بناء نماذج رياضية (إحصائية غالبًا) من بصمات الإشارات المقابلة لأنماط عينات التدريب، ومن ثَمَّ يجري حفظ هذه النماذج بكفاءة في قاعدة بيانات لاستدعائها في إحدى آليات التصنيف التي تقرر أي الأنماط هي الأقرب للإشارات المقابلة للأنماط المدخَلة المراد التعرف عليها.

ويعكسُ هذا الإطارُ نظرية «التعلم الحاسوبي» (Machine Learning) التي تُطبَّق كأنسب ما يكون من أجل مقارَبة المسائل التي لا يُعْرَف لها أو حيث يتعذر الحصول لها على «حلولٍ بقواعد مُحْكَمة» (Closed-Form Solutions)، وتنبثق مختلف أساليب التعلم الحاسوبي من مبدأ إمكانية التعلم عبر تكرار التعرض الغزير لكل من الأمثلة الصحيحة والأمثلة الخاطئة أو عبر تكرار التعرض الغزير للأسئلة وأجوبتها حول جزئيات المسألة المطلوب مقاربتُها؛ فالطفل مثلاً قد يجيد قراءة النص المكتوب بلغته الأم قبل أن يتعلم أسس وقواعد اللغة وذلك عبر تقليد الأكبر سنًّا وعبر محاولات الصواب والخطأ مع تصويب الأخطاء حيث يتحسن الأداء مع معاودَة ذلك كله مرارًا وتكرارًا.



الشَّكل ١-١٧: الهيكل الوظيفي العام لأنظمة التعرف على الأنماط وفق نظرية التعلم الحاسوبي.

ويرتكز بصفة عامة تفعيل تلك الأساليب رياضيًّا وحاسوبيًّا لهذا المبدأ على استقراء السياق الاحتمالي للكلمات وحروفها (المناظِرة للأنماط التي ندرسها في هذا الفصل) عِوَضًا عن سياقها اللغوي بُغْيَةَ الوصول إلى حساب الاحتمال الرياضي لوقوع كل تعرُّفٍ ممكنٍ للكلمة بين ما يسبقه وما يلحق به من كلمات ومن ثَمَّ ترجيحِ التعرف صاحب أعلى احتمال رياضي، وتستلزم عملية إجراء الحساب هذه تكوين نموذج احتمال رياضي يحاكي كل تتابُعات الوحدات اللغوية كما تحدث في الواقع الحقيقي لاستخدام اللغة.

وأيًّا ما كانت الأساليب الرياضية والحاسوبية المختارة لبناء مثل هذا النموذج الاحتمالي، فلا بُدَّ لها من بيانات تجريبية غزيرة تملأ وعاءً واسعًا من بصمات الأنماط (الأسئلة) بالتوازي مع سلاسل أكواد الأنماط المقابلة لهذه البصمات (أجوبتها) وذلك مثلما يقتضي التعلم الحاسوبي الموجَّه، ويُطْلَق على عملية تشغيل هذه الأساليب الرياضية على محتويات هذا الوعاء من أجل بناء النموذج الاحتمالي اسْمَ «التدريب» (Training) كما يُطْلَق على محتويات هذا الوعاء اسم «مَوْرِد لُغَويّ» (Language Resource)، وبطبيعة الحال فإن النموذج الاحتمالي سوف يحمل الخصائص الإحصائية للمورد اللغوي الذي تَكَوَّنَ منه، والذي عليه بدوره أن يحمل الخصائص الإحصائية والسياقية للأنماط الذي يعبر عنها كما سنتعرض إليه في القسم الخامس من هذا المبحث.

وتُعَدُّ المكوِّنات الثلاث التي تميز أيَّ نظام بعَيْنه للتعرف على الأنماط عما سواه ضمن

هذا الإطار هي تلك التي تقوم باستخلاص بصمة الإشارة، وببناء نماذج رياضية للأنماط، وبالتصنيف الآلي للبصمات المستخلصة من أجل تقرير أقرب الأنماط المقابلة لها.

وسوف يُلْمِح كُلٌّ من الأقسام الثلاثة التالية باختصار إلى أنجح المقارَبات المستخدَمة[1] لتنفيذ كُلٍّ من هذه المكونات الثلاث في كل حالة من الحالات الثلاث الرئيسية لتقنيات التعرف على الكتابة كما صنفها القسم الأول من المبحث السابق.

٢- أساليب التعرف على الكتابة العربية بتعقب خط اليد

تحققت أفضلُ النتائج خلالَ العِقدين الماضيين في التعامل مع هذه المسألة عبر المعالَجات الرياضية وفق إحدى منهجيتين [١، ١٢]؛ منهجية «نماذج ماركوف المخفِيَّة» (Hidden Markov Models: HMM) وهي ما سنعرِّج عليه في القسم التالي، وكذلك منهجية «الشبكات العصبية الاصطناعية» (Artificial Neural Networks) التي تستلهم آليات عمل الخلايا العصبية الحقيقية من حيث الاستثارة والاستجابة وارتباطاتها معًا في شبكات كثيفة حيث تقوم كلٌّ من هذه الشبكات العصبية الاصطناعية بدور الدالَّة (أي العلاقة) الرياضية التقريبية التي تربط بين المدْخَلات (وهي هنا الحروف) إلى الظاهرة/ النظام/ العملية قَيْدَ الدراسة (وهي هنا عملية الكتابة العربية بتعقب خط اليد) وبين الملاحظات المرصودة نتيجةَ هذه العملية (وهي الكتابة المرقمَنة - كما ورد توصيفُها في القسم الأول من المبحث السابق - و/ أو كمياتٌ مشتقةٌ منها)[2].

ويُعْرَف العديدُ من أنواع هذه الشبكات العصبية الاصطناعية التي تلائم كلٌّ منها طائفةً معينة من المسائل من تلك التي يتعــذر الحصول لهــا علــى «حلول رياضية بقــوانيــن

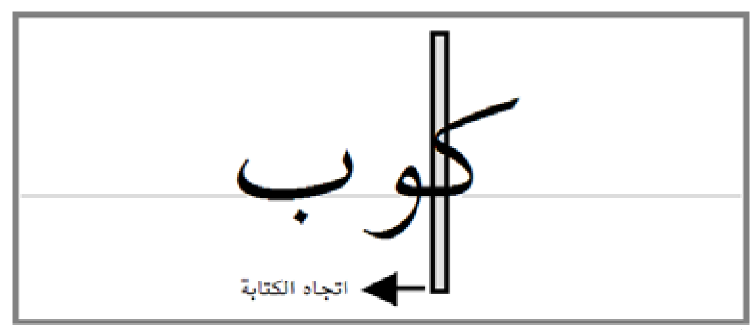
١- يتطلب استيفاء كل مقارَبة من مثل تلك المقارَبات تخصيص عشرات الصفحات لتفاصيل فنية يتطلب شرحها الكثير من الرياضيات المتقدمة، وهو ما لا يتسع له المجال في هذا المقام، وسوف نحيل القارئ المهتم بهذه التفاصيل إلى المراجع المختصة بذلك.

٢- نحيل القارئ المهتم بالتفاصيل الرياضية والهندسية بمحاكاة الشبكات العصبية واستخداماتها في التعلم الحاسوبي إلى الورقة التعليمية الآتية:
Jain, A.K., Jianchang Mao, Mohiuddin, K.M., Artificial Neural Networks: A Tutorial, IEEE Computer Magazine, pp. 31–44, Mar. 1996.
http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=485891

مُحْكَمة» (Closed Form Solutions)، ويمكـــن استخدامُ هذه الشبكات للربط في كـــلا الاتجـــاهين بين المدخــــلات وما تـــؤدي إليه من ملاحظـــات مرصـــودة - أي الحصـــول على إحدى المجموعتين بمعلومية المجموعة الأخرى - وبالطبع فإن الربط العكسي هو الأداة المستخدَمة للتعرف على حروف الكتابة المقابلة لما لدينا من منحنَياتٍ مرقْمَنةٍ تمثِّل تعقب خط اليد أثناء الكتابة. وبطبيعة الحال فإن ذلك كله يتطلب عملية تدريب للشبكة المستخدَمة على كميات كبيرة من المدخَلات بالتوازي مع ما يناظرها من ملاحظات مرصودة وهو ما سنعرض إليه في القسم الخامس من هذا المبحث.

٣- أساليب التعرف على الكتابة العربية المطبوعة

بتتبع الأبحاث الجادة التي جرت خلال العقدين الماضيين - العقد الأخير من القرن العشرين والعقد الأول من القرن الحادي والعشرين - فإن تلك التي أثمرت أفضل النتائج في التعامل مع مسألة التعرف الآلي على الكتابة العربية المطبوعة [٥، ٧، ١١] اشتركت في أنها تستخدم المعالَجات الرياضية القائمة على «نماذج مارْكُوفْ المخفِيَّة» (Hidden Markov Models) التي تسعى رغم اختلاف تنويعاتها إلى استنباط «تسلسل المدخَلات» (Inputs Sequence) الأرجح احتماليًّا (المجهولة لدينا) الذي يؤدي عبر الظاهرة/ النظام/ العملية قَيْدَ الدراسة (عملية طباعة الكتابة في حالتنا هذه) إلى «تسلسل الملاحظات» (Observations Sequence) (المعلومة لدينا)[١].



الشَّكل ١-١٨: النافذة المنزلقة في اتجاه الكتابة على سطور الصورة الممسوحة ضوئيًّا.

١- نحيل القارئ المهتم بالتفاصيل الرياضية والهندسية لنماذج ماركوف المخفِيَّة إلى الورقة التعليمية الآتية:
Rabiner, L.R., A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, Volume 77 - No. 2, pp. 257–286, Feb. 1989.

ويناظر تسلسلُ المدخَلات في حالتنا هذه ترميزاتِ («أكواد») حروفَ الكلمات المطبوعة المراد التعرف عليها، أما تسلسلُ الملاحظات فهو تسلسلُ «متَّجَهاتٍ» (Vectors) يصوغ كلٌّ منها رياضيًّا بعضَ خصائص المحتوى الرسومي داخل «نافذة منزلقة» (Sliding Window) على سطور الصورة الممسوحة ضوئيًّا والمرقْمَنة الجاري التعرف عليها كما في الشكل رقم (١-١٨) أعلاه.

وبينما تمتاز المعالجات المرتكزة على نماذج ماركوف المخفِيَّة بقدرتها على فَصْل جرافيمات الكتابة العربية المتصلة مع التعرف على هذه الجرافيمات في آن واحد، فإن تصميم متجه خصائص التعرف كان -ولا يزال- اجتهادًا مفتوحًا للباحثين تتفاوت صُوَرُهُ بين تصميمات بسيطة مثل عدد البُقَع المتصلة في النافذة الرأسية المنزلقة فوق «سطر الكتابة المعياري» (Baseline) [1] مع عدد البقع المتصلة تَحْتَه، وبين تصميمات أعمق رياضيًّا مثل حساب «العُزُوم القياسية اللامتغيرة» (Invariant Moments) للمحتوى الرسومي للنافذة. [٥، ٧، ٨، ١١] وأيًّا كان التصميمُ المختار فإن هناك معايير للحكم على مدى جودة خصائص التعرف على الكتابة نُجْمِلها فيما يلي:

١. ضرورة أن يكون عدد المركِّبات في متجه خصائص التعرف محدودًا.

٢. ضرورة أن يكون هذا العدد المحدود من المركِّبات في متجه خصائص التعرف ثابتًا.

٣. الحرص على أن يكون عدد المركِّبات في متجه خصائص التعرف أقل ما يمكن.

٤. الحرص على قابلية حساب كُلِّ مركِّبات متجه خصائص التعرف بكفاءة.

٥. الحرص على أن تبقى تسلسلات متجهات خصائص التعرف المناظِرة لنفس جرافيمات الكتابة ثابتةً قدر الإمكان برغم أي اختلافات في أشكال هذه الجرافيمات من حيث الأبناط والأحجام، أو بسبب بعض التشوهات الأخرى مثل الميل المحدود في سطور الكتابة أثناء مسحها ضوئيًّا أو احتواء الكتابة على بعض الشوشرة المحدودة.

١- وهو السطر الأفقي الرمادي في الشكل (١-١٨) أعلاه.

وبينما يؤدي الفشل في تحقيق الشرطين الأول والثاني مباشرةً إلى تعذر تطبيق أية معالَجات قائمة على نماذج ماركوف المخفِيَّة - وكذلك أية أساليب أخرى للتعلم الحاسوبي على وجه العموم - فإن تحقق الشرطين الثالث والرابع ضروري لإنتاج أنظمة واقعية للتعرف على الكتابة المطبوعة ذات تكلفة حاسوبية معقولة من حيث زمن التشغيل ومساحات التخزين، أما تحقق الشرط الخامس فهو مفتاح التمكن من التعرف على الكتابة العربية المطبوعة رغم تعدد أبناط وأحجام الكتابة.

٤- أساليب التعرف على الكتابة العربية المخطوطة يدويًّا

بسبب زيادة التنوع وعدم الانتظام في رسم الكتابة المخطوطة يدويًّا بالمقارَنة مع الكتابة المطبوعة، فإن هوامش الخطأ في أداء أنظمة التعرف على الكتابة العربية المخطوطة يدويًّا هي بطبيعة الحال أكبرُ كثيرًا مقارَنةً بهوامش الخطأ في أداء أنظمة التعرف على الكتابة العربية المطبوعة. ومع ذلك، فإن أفضل المحاولات التي جرت خلال العقدين الماضيين للتعرف على الكتابة العربية المخطوطة يدويًّا اشتركت في ارتكازها على المعالَجات الرياضية القائمة على نماذج مارْكُوفْ المخفِيَّة بأسلوب مشابه لما جرى عرضُهُ في القسم السابق. [١٠، ١٣، ١٤].

٥- بناء الموارِد اللغوية لتدريب وتقويم أنظمة التعرف على الكتابة العربية

يتضح من الأقسام السابقة في هذا الفصل أن الأساليب السائدة والأكثر نجاعةً حتى وقتنا هذا للتعامل مع مسائل التعرف الآلي على الكتابة العربية بتنويعاتها المختلفة هي أساليب التعلم الحاسوبي بأصنافها المتعددة، وكما جاء في ختام القسم الثامن من باب «التحليل الصرفي الآلي لمفردات اللغة العربية» من هذا الكتاب فإن كل هذه الأساليب يستلزم تفعيلُها تكوينَ نماذج احتمالية رياضية تحاكي كل تتابعات جرافيمات الكتابة العربية كما تحدث في الواقع الحقيقي لنصوص اللغة وذلك بُغْيَةَ المساعدة في اختيار تتابع الجرافيمات الذي يحقق أعلى احتمال رياضي في سياق التعرف على أشكالها المخطوطة أو المطبوعة.

ومن أجل بناء تلك النماذج الرياضية فلا بُدَّ من مادة لغوية تملأ وعاءً واسعًا من الأسئلة (وهي في حالتنا هذه الكتابةُ المرقْمَنة مطبوعةً كانت أم مخطوطةً) بالتوازي مع أجوبتها (النص العربي المناظِر لهذه الكتابة المرقمنة) لتغذية أساليب «التعلم الحاسوبي الموجَّه» (Supervised Machine Learning)، ويُطْلَق على عملية تشغيل هذه الأساليب الرياضية على المادة اللغوية من أجل بناء النماذج الاحتمالية الرياضية اسم «التدريب» (Training) كما يُطْلَق على الوعاء الممتلئ بتلك المادة اللغوية اسم «المورد اللغوي» (Language Resource)، وبطبيعة الحال فإن النماذج الاحتمالية سوف تحمل الخصائص الإحصائية للمورد اللغوي الذي تَكَوَّنَت منه، والذي عليه بدوره أن يحمل خصائص النصوص العربية المكتوبة الذي يعبر عنها.

ولذلك فإن بناء هذه الموارِد اللغوية يشترط إضافةً إلى الدقة العالية أن يجري التمثيلُ بصورة متوازنة لمختلف أبعاد التنوع في الكتابة مثل؛ أنماط الطباعة وأحجامها، أو الخطوط اليدوية للكُتَّاب ... إلخ.

ففي حالة أنظمة التعرف على الكتابة العربية بتعقب خط اليد يتكون المورد اللغوي من الكتابة العربية المرقْمَنة -كما جاء توصيفُه في القسم الأول من هذا الفصل- بالتوازي مع النص العربي الصحيح المقابل له، وتكون المحاذاة بين هاتين المركَّبَتَيْن على مستويين؛ أولاهما على مستوى «الضربات المتصلة بالقلم» (Strokes) (ويناظر ذلك تقسيمُ العبارة السابقة التي تحتها خط إلى: ا ، لضر ، با ، ت ، ا ، لمتصلة ، با ، لقلم)، وثانيهما على مستوى الكلمات العربية الكاملة.

أما في حالة أنظمة التعرف على الكتابة العربية المطبوعة فإن المورد اللغوي يتكون من رَقْمنة صفحات الكتابة العربية الممسوحة ضوئيًّا - كما جاء توصيفُه في القسم الأول من هذا الفصل - بالتوازي مع النص العربي الصحيح المقابل له، وتكون المحاذاة بين هاتين المركَّبَتَيْن على مستوى الكلمات العربية الكاملة وكذلك على مستوى السطور في كل صفحة، ولذلك الغرض تُسْتَخْدَم أدواتٌ براجميةٌ للتقسيم الآلي لكل صفحة مطبوعة إلى كلمات وسطور كما هو مبين في الشكل الموضَّح أدناه، وعلى اللغويين القائمين على بناء هذه الذخيرة اللغوية التحققُ من صحة هذا التقسيم الآلي وتصويبه عند اللزوم.



الشَّكل ١-١٩: نموذجٌ على نواتج تعيين حدود الكلمات العربية الكاملة وحدود السطور بصورة آلية اعتمادًا على التوزيع التكراري لنقاط الصفحات المرقْمَنة.

ويتشابه كثيرًا بناءُ الموارد اللغوية لأنظمة التعرف على الكتابة العربية المخطوطة يدويًّا مع تلك المتعلقة بأنظمة التعرف على الكتابة العربية المطبوعة مع ملاحظة أن عمليات التقسيم الآلي للصفحات إلى كلمات وسطور ثم تصويب هذا التقسيم بل وتفسير النصوص في الصفحات المخطوطة يدويًّا تستغرق من اللغويين عملاً ووقتًا أكبر.

يُشكِّل بناءُ الموارِد اللغوية اللازمة لتدريب أنظمة التعرف على الكتابة العربية بأنواعها الثلاثة؛ «بتعقب خط اليد» و«المطبوعة» و«المخطوطة يدويًّا» نسبةً كبيرةً من أنشطة واستثمارات البحث والتطوير في هذه التقنيات، يُقدِّرُها بعضهم بحوالي خمسين في المئة، وبسبب ضخامة الحجم وارتفاع دقة العنونة المطلوبتين أثناءَ بناء كلٍّ من هذه الموارِد فإن هذا النشاط يستدعي توظيف عدد من اللغويين العرب المؤهَّلين حاسوبيًّا لإنجازه على نحو مُرْضٍ مما نَعُدُّه أحد الأهداف التي نرجو أن يحققها تأليف هذا الفصل من هذا الكتاب.

ببليوجرافيا مرجعيَّة

المبحث الأوَّل

١. الحَمَد (غانِم قَدُّوريّ): رَسْمُ المصْحَفِ؛ دِراسةٌ لُغَوِيَّةٌ تاريخيَّةٌ، ط١، بَغْداد، ١٩٨٢.

٢. الدَّالي (عبد العزيز): الخِطاطة «الكِتابة العَرَبيَّة»، مَكتبة الخانْجِي، مِصْر، ١٩٩٨م.

٣. دُرْمان (مُصْطَفَى أَغْوَر): فَنُّ الخَطِّ العَرَبِيّ؛ مَوْلِدُهُ وتَطَوُّرُهُ حَتَّى العَصْرِ الحاضِر، تَرْجَمة: صالِح سِعْداوِيّ، ط١، إستانبول، ١٩٩٠م.

٤. عَفيفيّ (فَوْزيّ سالِم): نَشْأة وتَطوُّر الكِتابة الخَطِّيَّة العَرَبيَّة ودَوْرُها الثَّقافيّ والاجتماعيّ، ط١، وِكالةُ المطبوعات، الكُوَيْت.

5. Attia, M., Arabic Orthography vs. Arabic OCR; Rich Heritage Chal-lenging a Much Needed Technology: http://multilingual.com/issueDetail.php?issue=68, Multilingual Computing & Technology magazine: www.Multilingual.com, USA, Dec. 2004.

المبحث الثَّاني

1. Abdou, S., Fahmy, A., Hosney, I., Mostafa, I., Artificial Tutor for Arabic Handwriting Training, The Proceedings of the 2nd International Conference on Arabic Language Resources and Tools, ISBN: 2-9517408-5-9; http://www.medar.info/conference_all/2009/index.php, Cairo-Egypt, Apr. 2009.

2. Al-Badr, B., Mahmoud, S.A., Survey and Bibliography of Arabic Optical Text Recognition, Elsevier Science, Signal Processing 41, pp. 49-77, 1995.

3. Attia, M., Rashwan, M., El-Mahallawy, M., Autonomously Normalized Horizontal Differentials as Features for HMM-Based Omni Font-Written OCR Systems for Cursively Scripted Languages, on IEEE Xplore at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5478619: The Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSI-

PA09) http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp? punumber=5472904, pp. 185 to 190, Kuala Lumpur - Malaysia http://www.SP.ieeeMalaysia.org/ICSIPA09, Nov. 2009.

4. Attia, M., Arabic Orthography vs. Arabic OCR; Rich Heritage Challenging a Much Needed Technology http://multilingual.com/ issueDetail.php?issue=68, Multilingual Computing & Technology magazine: www.Multilingual.com, USA, Dec. 2004.

5. Callan, J., Kantor, P., Grossman, D., Information Retrieval and OCR: from Converting Content to Grasping Meaning, SIGIR conference, 2003.

6. Duda, R.O., Hart, P.E., Stork, D.G., Pattern Classification, 2nd edition, John Wiley & Sons Inc., 2001.

7. Gouda, A.M., Arabic Handwritten Connected Character Recognition, PhD thesis, Dept. of Electronics and Electrical Communications, Faculty of Engineering, Cairo University, Nov. 2004.

8. Govindan, V.K., Shivaprasad, A.P., Character Recognition; a Review, Pattern Recognition, Vol. 23, No. 7, pp. 671–683, 1990.

9. Jumari, K., Ali, M.A., A Survey and Comparative Evaluation of Selected Off-Line Arabic Handwritten Character Recognition Systems, Jurnal Teknologi - Universiti Teknologi - Malaysia, 36(E) June 2002: 1–18.

10. Lorigo, L.M., Govindaraju, V., Offline Arabic Handwriting Recognition: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 5, pp. 712 to 724, May 2008.

11. St. Clair, G., Million Book Project vs. Google™ Print, Journal of Zhejiang University SCIENCE (JZUS) http://www.zju.edu.cn/jzus-China, ISSN 1009-3095, 2005-6A(11): pp. 1195-1200.

12. Abdelazim, H. Y., Malek, M., System and Method for Onscreen Text Recognition for Mobile Devices, US Patent and Trademark Office, Application No. US 12/196925 filed on Aug. 22nd 2008, Publication No. US 2009/0055778 A1 published on Feb. 26th 2009.

13. Al-Badr, B., Mahmoud, S.A., Survey and Bibliography of Arabic Optical Text Recognition, Elsevier Science, Signal Processing 41, pp. 49-77, 1995.

14. Al-Ma'adeed, S., Elliman, D., Higgins, C.A., A Data Base for Arabic Handwritten Text Recognition Research, IEEE Proceedings of Frontiers in Handwriting Recognition, Nov. 2002.

15. Al-Ohali, Y., Cheriet, M., Suen, C., Databases for Recognition of Handwritten Arabic Cheques, In the Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, Amsterdam, pp. 601-606, 11-13 Sept. 2000.

16. Bazzi, I., Schwartz, R., Makhoul, J., An Omnifont Open-Vocabulary OCR System for English and Arabic, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 6, June 1999.

17. Duda, R.O., Hart, P.E.; Stork, D.G., Pattern Classification, 2nd edition, John Wiley & Sons Inc., 2001.

18. El-Mahallawy, M.S.M., A Large Scale HMM-based Omni Front-Written OCR System for Cursive Scripts, PhD thesis, Dept. of Electronics and Electrical Communications, Faculty of Engineering, Cairo Univ., Apr. 2008.

19. Gouda, A.M., Arabic Handwritten Connected Character Recognition, PhD thesis, Dept. of Electronics and Electrical Communications, Faculty of Engineering, Cairo University, Nov. 2004.

20. Govindan, V.K., Shivaprasad, A.P., Character Recognition; a Review, Pattern Recognition, Vol. 23, No. 7, pp. 671–683, 1990.

21. Jumari, K., Ali, M. A., A Survey and Comparative Evaluation of Selected Off-Line Arabic Handwritten Character Recognition Systems, Jurnal Teknologi - Universiti Teknologi - Malaysia, 36(E) June 2002: 1–18.

22. Khorsheed, M. S., Offline Recognition of Omnifont Arabic Text Using the HMM Tool Kit (HTK), Pattern Recognition Letters – Elsevier, Vol. 28, pp. 1563 – 1571, 2007.

23. Korb, J., Survey of Existing OCR Practices and Recommendations for More Efficient Work, TELplus EC funded project; downloadable at: https://www.zotero.org/mchristy/items/itemKey/S5VCTFRT, July 2008.

24. Lorigo, L.M., Govindaraju, V., Offline Arabic Handwriting Recognition: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 5, pp. 712-724, May 2008.

25. Mohamed, M., Gader, P., Handwritten Word Recognition Using Segmentation-Free Hidden Markov Modeling and Segmenta-tion-Based Dynamic Programming Techniques, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 5, pp.548–554, 1996.

26. Strassel, S.M., Linguistic Resources for Arabic Handwriting Recognition, The Proceedings of the 2nd International Conference on Arabic Language Resources and Tools, ISBN: 2-9517408-5-9, Cairo - Egypt http://www.medar.info/conference_all/2009/index.php, Apr. 2009.

الفصل الثَّاني
مُعالَجة النَّصّ العربيّ المنطوق

د. مُحمَّد عفيفي

المبحث الأوَّل: التَّعرُّف الآليّ على الكلام.

المبحث الثَّانِي: نظم تحويل النصّ إلى كلام.

المبحث الثَّالث: نُظُم التَّعَرُّف على اللُّغة والمتكَلِّم.

المبحث الأوَّل
التَّعرُّف الآليّ على الكلام

١ – مقدمة.

٢ – مكونات نظم التعرف على الكلام.

٣– ملخص لبناء نظام للتعرف على الكلام.

٤ – التعرف على الكلام في اللُّغة العربيَّة.

١ – مقدمة

نظم التعرف الآليّ على الكلام لها الكثير من التطبيقات المهمة في مختلف المجالات. ومن المعروف أن هذه النظم – وخاصة المعاصرة منها – تعتمد على مجموعة من النماذج الإحصائية التي تعبر عن الأصوات المختلفة في اللغة التي يراد التعرف عليها؛ وبما أن الكلام – كما تبين من الفصول السابقة – له بناء زمني وطيفي، بمعنى أن كل صوت من الأصوات هو عبارة عن تسلسل زمني لبعض المتجهات الطيفية، فإن نماذج ماركوف المخفيَّة (Hidden Markov Models –HMM) تعتبر من أهم (إن لم يكن أهم) النماذج التي تُستخدم لبناء نماذج الأصوات في نظم التعرف على الكلام المعاصرة.

إن أسس نظم التعرف على الكلام المعاصرة – والقائمة على نماذج ماركوف المخفِيَّة – ظهرت في سبعينيَّات القرن الماضي في جامعة «كارنيجي ميلون» (CMU) الأمريكية، وكذلك في شركة «آي بي إم» (IBM)؛ وفي هذه الأثناء كانت النماذج قائمة على استخدام نماذج الكثافة المتقطعة (discrete density)، والتي سرعان ما تطورت معامل «بيل» (Bell labs) إلى استخدام ما يُعرَف بـ «الكثافة المتصلة» (continuous density). وكانت الأبحاث في هذا الوقت تنصب على النظم ذات الكلمات المنفصلة من متكلم واحد، أو النظم التي تتعرف على عدد محدود من الكلمات (الأرقام على سبيل المثال). ومع تسعينيَّات القرن بدأ الاهتمام بنظم الكلام المتصل التي لا تعتمد على المتكلم. وكان الدافع الرئيسي لهذه الأبحاث سلسلة من البرامج التي كانت تمولها وكالة مشروعات الأبحاث الدفاعية المتقدمة Defense Advanced Research Projects Agency (DARPA)، والتي عملت على تحسين نسبة التعرف بالنسبة لأنواع مختلفة من الكلام المتدرجة الصعوبة. ومن المناسب في هذا السياق أن نذكُرَ أن اهتمام هذه الهيئة بتمويل أبحاث اللغة العربية بدأ في خلال العام ٢٠٠٣م، واستمر في عدة مشروعات بحثية أدت إلى تحسن كبير في دقة التعرف على اللغة العربية؛ وسنتعرَّض لبعض هذه الأمثلة في هذا المبحث.

وتحقيقًا للغاية من هذا المبحث سنبدأ بشرح المكونات الأساسية لنظم التعرف على الكلام القائمة على نماذج ماركوف المخفِيَّة – وبالتحديد سنقوم بشرح النظم التي تعتمد على نماذج (جاوس) البسيطة؛ كما سنتطرَّق إلى كيفية حساب معاملات هذه النماذج آليًّا من خلال إشارات الكلام. ويجدر القول في هذا السياق أن إمكانية حساب هذه

المعاملات آليًّا تُعَدُّ أحد أهم أسرار النجاح لنماذج ماركوف المخفِيَّة. ومع طُرُق حساب المعاملات تأتي مناقشة دالة الهدف (Objective function) التي يتم الاستعانة بها أثناء تدريب النماذج؛ كما سنتعرَّضُ لشرح بعض التطبيقات العملية لهذه النظم وخصوصًا تلك الَّتي تُعنى باللغة العربية.

٢- مكونات نظم التعرف على الكلام

يتم تحويل إشارة الكلام الداخلة إلى سلسلة من المتجهات الأكوستيكية، يطلق عليها - عادةً - «مُتَّجهات السِّمات» (feature vectors)؛ وتسمى هذه العملية «استنباط السِّمات» (feature extraction)، ثم يقوم محرك البحث بمحاولة إيجاد سلسلة الكلمات المناظرة، بحيث تكون:

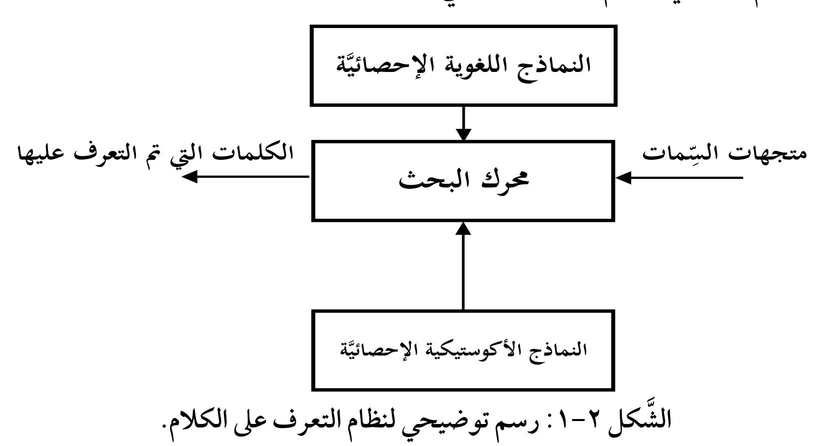
$$\hat{w} = \underset{w}{argmax}\ p\,(w|y) \qquad (١)$$

ما يتم تبسيط المعادلة رقم (١) باستخدام قاعدة Bayes إلى:

$$\hat{w} = \underset{w}{argmax}\ p\,(y|w)\ p\,(w) \qquad (٢)$$

تُشير $\hat{w}$ إلى أفضل سلسلة من الكلمات، وتُشير w إلى أية سلسلة من الكلمات.

ونحسب الكمية $p\,(y|w)$ استخدام النموذج الأكوستيكي الإحصائي (HMM) والكمِّيَّة باستخدام النموذج اللغوي الإحصائي. والشكل (٢-١) يبين نظام للتعرف على الكلام، والذي سيتم تفصيله فيما يلي:



الشَّكل ٢-١: رسم توضيحي لنظام التعرف على الكلام.

النموذج الأكوستيكي: كما هو معروف فإن أصغر وحدة صوتية هي الفونيم. فعلى سبيل المثال، كلمة كَتَبَ تتكون كالتالي /k /a /t /a /b /a/ . وقد ناقشنا الوحدات الصوتية للغة العربية في الفصول السابقة. لذلك فإن أي كلمة يمكن تمثيلها عن طريق ربط الفونيمات المكونة لها كما في المعجم النطقي.

ويمكن اعتبار النموذج الناتج هو النموذج الأكوستيكي لهذه الكلمة؛ ومن الجدير بالذكر أن معاملات نماذج الفون المكونة لأي كلمة يتم حسابها باستخدام ما يسمى «بيانات التدريب» والتي تتكون من إشارات الكلام والكلمات المناظرة. ويتم هذا التدريب في مرحلة بناء النظام[1].

النموذج اللغوي: تتكون النماذج اللغوية - غالبًا - عن نماذج (N-gram)، وهذا النموذج يعطي احتمال ظهور كلمة ما مشروطًا على الكلمات N-١ السابقة. فعلى سبيل المثال: فإن نموذج (2-gram) أو (bigram) يحسب احتمال ظهور كلمة الولد بعد كتب (الولد|كتب)p، وباستخدام هذه النماذج يمكن حساب احتمال أي جملة من الجمل من الناحية اللغوية، وسيتم تفصيل هذه النماذج لاحقًا.

وبعد حساب النموذجين -الأكوستيكي واللغوي - فإن برنامج فاكّ الشَّفرة (decoder) يمكنه حساب أرجح سلسلة من الكلمات المناظرة لإشارات الكلام. ومن الجدير بالذكر أنه في نظم التعرف ذات العدد الكبير من الكلمات فإن عمليات البحث تتطلب قدرات حسابية عالية، ولذلك يلزم تطوير محركات بحث ذات كفاءة عالية؛ وعليه فإن إجراء البحث على عدة مراحل باستخدام «تشبيكة» (lattice) يعتبر من الطرق الواسعة الانتشار، وسيتم الحديث عن ذلك لاحقًا.

٢,١ - استنباط السِّمات (Feature extraction)

تعد خطوة استنباط السِّمات خطوة أساسية تهدف إلى استنتاج المعلومات المهمة في إشارات الكلام، وفي نفس الوقت تقلل من فقد المعلومات. وبالإضافة إلى استنباط صورة مبسطة لإشارات الكلام فإن خطوة استنباط السِّمات تهدف - كذلك - إلى

١- يُرجَى ملاحظة أنَّنا - في هذا الفصل - نستخدم كلمتَي «فون» و «فونيم» بدون تمييز؛ مع التأكيد علي وجود بعض الفروق التقنية بين الكلمتين.

الحصول على متجهات سِمات تناسب إلى حد كبير فروض النماذج الأكوستيكيَّة. ومن المتعارف عليه أن متجهات السِّمات يتم استخلاصها كل عشرة مِيلِّي ثانية من نوافذ طولها ٢٥ مِيلِّي ثانية. ومن أشهر الطرق المستخدمة: المعاملات (MFCC – Mel frequency cepstral coefficients)، ويتم حساب هذه المعاملات عن طريق تطبيق «تحويل جيب التَّمام المتَقَطِّع» (Discrete cosine transform) على اللوغاريتم الطيفي، ويتم الاستعانة بمقياس Mel اللاخطي لمضاهاة ترددات الأذن. وبالإضافة إلى MFCC فإن معاملات التنبؤ الخطي الإدراكية (Perceptual Linear Prediction – PLP) تقوم بتحويل معاملات التنبؤ الخطي المعروفة إلى معاملات (كِبِسْتِرَم) بعد تطبيق التحويلات اللاخطية المرتبطة بالإدراك. ومن المعروف أن مُعاملات MFCC ومُعاملات PLP تُعطي نتائج متقاربة في نظم التعرف على الكلام، مع بعض التَّميُّز النسبيّ لمعاملات PLP في أوساط الضوضاء. وبالإضافة لاستخدام المعاملات التي تعبر عن الطيف لإشارات الكلام فإنه من المعروف أن إضافة المشتقة الأولى والمشتقة الثانية لهذه المعاملات يستخدم على نطاق واسع في نظم التعرف على الكلام. فعلى سبيل المثال، إذا كان التمثيل الطيفي يتكون من ١٣ معاملاً من معاملات MFCC فإن المتجه النهائي يضيف المشتقة الأولى والثانية لهذه المعاملات ليصبح طول المتجه هو ٣*١٣=٣٩.

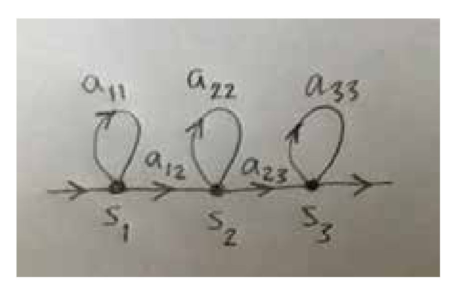
٢,٢- النماذج الأكوستيكية ونماذج ماركوف المخفِيَّة

كما ذُكِرَ في المقدمة، فإنه يمكن تمثيل الكلمة كسلسلة من الفونات وتسمى هذه السلسلة بالتمثيل الصوتي للكلمة ويمكن الحصول عليها باستخدام قاموس صوتي وفي العموم يمكن أن يكون للكلمة الواحدة تمثيلات صوتية متعددة تناظر طرق نطقها.

وأبرز مثال على ذلك في اللغة العربية هو التصريفات المختلفة للكلمة. فعلى سبيل المثال، فإن كلمة **كتب** يمكن نطقها **كَتَبَ** أو **كُتُب**.

ويتكون كل تمثيل صوتي عن طريق ربط عدد من الفونيمات. فعلى سبيل المثال فإن التمثيل الصوتي كَتَبَ للكلمة كتب يمكن تكوينه عن طريق ربط الفونيمات المكونة للتمثيل الصوتي ka-ta-ba.

ويتم تمثيل كل فونيم بما يسمى نماذج ماركوف المخفية كما هو موضح في الشكل (٢-٢).



الشَّكل ٢-٢: تمثيل الفونيمات باستخدام نماذج ماركوف المخفِيَّة.

ويتكون النموذج مما يلي:

- عدد من الحالات (states)، ثلاثة في الشكل.
- عدد من الانتقالات (transitions). ولكل انتقال ما يعرف باحتمال الانتقال. فمثلا المعامل a_{12} يعبر عن احتمال الانتقالات من الحالة رقم ١ إلى الحالة رقم ٢. ويكون مجموع احتمالات الانتقالات التي تخرج من حالة ما مساويا لواحد. ويُبَيِّن الشكل (٢-٢) نموذج ماركوف من اليسار إلى اليمين (left-to- right) الواسع الانتشار في نظم التعرف على الكلام.
- نماذج المخرجات (output distributions) لنماذج (جاوس) البسيطة الواسعة الانتشار في نظم التعرف على الكلام. وفي هذه الحالة تكون معاملات النموذج هي متجه المتوسط (mean vector) ومصفوفة التَّبايُن (Covariance Matrix).

ودون الدخول في تفاصيل أكثر، فإننا إذا أعطينا متجهات السِّمات المناظرة لنموذج ما فإنه يمكن حساب معاملات الانتقال ومعاملات نماذج المخرجات باستخدام ما يسمى الخوارزم الأماميّ الخلفيّ (forward-backward). وكذلك إذا أعطينا نماذج ما ومجموعة من متجهات السِّمات، فإنه يمكن حساب احتمال حدوث هذه المتجهات من النماذج.

ومما سبق فإنه إذا أعطينا حزمة من جمل التدريب فإنه يمكن أوتوماتيكيا حساب معاملات الفونيمات المكونة لهذه الجمل. ومن المعروف أن الفونات تتغير تغيرًا كبيرًا بتغير الفونات المجاورة، فعلى سبيل المثال: الفون (K) في كتب يختلف عن الفون (K) في سمك، وللتغلب على هذه الصعوبة يمكن استخدام نموذج لكل فون أخذًا في الاعتبار الفونات المجاورة من اليمين واليسار، وتسمى هذه النماذج «تراي فون» (Triphone). وعلى الرغم أن هذه النماذج تعبر بشكل أدق عن الفونات، فإن عددها يصبح كبيرا جدا. فعلى سبيل المثال إذا كان في لغة ما ٤٠ فون، فسينتج لدينا ٤٠ نموذج فوني أو أحادي الفونية و ٤٠×٤٠×٤٠ نموذجًا ثلاثي الفونية. وينتج عن هذا العدد صعوبة في حساب معاملات النماذج، ولذلك فإنه يتم ربط معاملات النماذج. ففي المثال السابق إذا وُجِدَ لدينا ٤٠×٤٠×٤٠ نموذج ثلاثي الفونية، أي حوالي ٦٤٠٠٠ نموذج. يمكن ربط هذه النماذج لعدد أقل - على سبيل المثال ٥٠٠٠ نموذج. ومع أن الربط فكرة جيدة وتعطي فرصة الموائمة بين عدد النماذج وحجم حزمة البيانات المتوفرة في التدريب، يبقى السؤال عن كيفية الربط أو المشاركة في المعاملات؛ ويتم هذا في أغلب الأحيان عن طريق شجرة القرار.

شجرة القرار (Decision Tree)

يتم ربط أو مشاركة النموذج عادة على مستوى الحالة وليس النموذج؛ وتستخدم شجرة القرار لتحديد أي الحالات يتم ربطها. ولكل حالة من كل فون q تكون هناك شجرة ثنائية، وتحتوي كل نواة في هذه الشجرة على سؤال عن جيران الفون.

ولربط النماذج المرتبطة بحال i في فون q فإن كل الحالات في النماذج المناظرة المشتقة من q يتم تجميعها عند جذر الشجرة. واعتمادًا على الإجابة على السؤال عند كل نواة يتم تقسيم الحالات إلى قسمين حتى الوصول إلى أوراق الشجرة.

وتتشارك كل الحالات الكائنة في ورقة من الأوراق في المعاملات. ويتم اختيار الأسئلة عند كل نواة من مجموعة مسبقة من الأسئلة. وعادة تكون هذه الأسئلة مرتبطة بجيران الفون.

فعلى سبيل المثال، تستخدم أسئلة مثل: هل على يمين الفون صوت متحرك؟

وهكذا يتم اختيار السؤال الذي يعطي أعلى زيادة في الاحتمالات عند الانقسام. في حالة استخدام نماذج جاوس الأحادية يمكن حساب الزيادة في الاحتمالية من الأعداد والمتوسطات دون الحاجة إلى البيانات الأصلية، ولذلك فإن عملية بناء الشجرة تتم بشكل سريع.

ملخص تدريب نماذج ماركوف:

- يتم اختيار نماذج أحادية الفون ابتدائية.
- يتم تدريب النماذج بتطبيق الخوارزمية الأمامية - الخلفية لعدد من المرات.
- يتم نسخ كل فون q إلى ثلاثي - فون x - q + y، والذي ظهرت باعتبارها حزمة في حزم التدريب. فعلى سبيل المثال: إذا ظهر الفون q (١٠٠٠) مرة فسيتم نسخه إلى ١٠٠٠ نموذج.

يتم تدريب نماذج ثلاثيَّة الفون باستخدام خطوة مشابهة، ولكن المشكلة أن بعض هذه النماذج تظهر لعدد صغير جدا من المرات.

- يتم تطبيق خوارزم شجرة القرار بمشاركة النسخ في عدد محدود من النماذج.
- يتم تدريب النماذج الناتجة بشكل مماثل.

٢,٣- النماذج اللغوية للنَّحو الإحصائي (N-gram)

يتم حساب احتمال سلسلة من الكلمات باستخدام نماذج النَّحو الإحصائِيّ N-gram حيث تكون N عادة في حدود ٢-٤.

ولتوضيح هذه النماذج يُفضَّلُ استخدام مثال. فلنأخذ الجملة «لقد قامت الثورة المصرية في الخامس والعشرين من يناير». وإذا أخذنا نظرة احتمالية للغة بما قد لا يروق لبعض اللغويين - ولكنَّه على أي حال ما نقوم به في نظم التعرف على الكلام- فإن احتمال كلمة الخامس في الجملة السابقة يكون:

$$P\,(\text{الخامس} \mid \text{لقد قامت الثورة في المصرية})$$

ويبسط نموذج النَّحو الإحصائي N-gram المثال السابق باعتبار الكلمات N-1 التي سبقت الكلمة. ففي الجملة السابقة إذا اعتبرنا نماذج النَّحو الثُّلاثيّ *3-gram*، فإننا

نكتب الاحتمال كما يلي:

(المصرية في | الخامس) P أو الثنائي جرام (في | الخامس) P

ومن الواضح أنَّه بزيادة قيمة N يتم تحسين القدرة التنبؤية للنموذج؛ ولكن على حساب:

- القدرة على حساب الاحتمالات.
- الكفاءة الحسابية للنماذج.

ويتم حساب نماذج (N-gram) ببساطة شديدة فعلى سبيل المثال فإنه احتمال (المصرية في | الخامس) P يتم حسابه ببساطة كعدد مرات حدوث «المصرية في الخامس» مقسوما على عدد مرات حدوث «المصرية في» في مدونة كبيره من النصوص. ورغم بساطة هذه الطريقة فإن المشكلة الواضحة بإن كثير من الاحتمالات ستكون صفرية لعدم مشاهدتها في مدونة النص. ويتم التغلب على هذه المشكلة بما يسمى أسلوب الخَصْم (discount) والتراجع (back-off) ومنها على سبيل المثال طريقة (Kneser-Ney).

٢,٤- محركات البحث

كما ذُكِرَ في مقدمة هذا الفصل فإن سلسلة أكثر الكلمات احتمالاً يمكن حسابها من متجهات السِّمات عن طريق البحث في أرجح سلسلة من الحالات التي يمكن أن تكون قد أنتجت سلسلة متجهات السِّمات طبقًا للمعادلة رقم ٢؛ ويمكن حل هذه المعادلة بكفاءة عن طريق البرمجة الديناميكية أو ما يسمى «خوارزم» (Viterbi)؛ وعند نهاية الجملة يمكن حساب أفضل احتمال. وإذا سجلنا أفضل اختيار عن كل لحظة، فإن بإمكاننا الرجوع والحصول على أفضل سلسلة من الكلمات.

وعلى الرغم من كفاءتها، فإن طريقة خوارزم (Viterbi) لا يمكن تطبيقها مباشرة في حالة وجود عدد كبير جدًّا من الكلمات، وخاصة مع نماذج لغوية طويلة. ولحل هذه المشكلة والحصول على محرك للبحث كفء، فقد ظهرت طرق عديدة، وسوف نعرض لهذه الطرق دون الدخول في تفاصيلها. ويمكن برمجة خوارزم (Viterbi) بكفاءة عن طريق تطبيق ما يسمى «التقليم» (Pruning) و«تمرير الرُّموز» (Token passing).

وتقوم الفكرة الأساسية على ألاَّ يتم حساب جميع الحالات الواقعة في فضاء البحث، ولكن يتم التركيز على الحالات التي تقع في نطاق الشعاع (beam width) من أرجح الحالات. وعلى الرغم من أن التقليم وتمرير الرُّموز يمكن أن يؤدي إلى تطبيق جيِّد لخوارزم (Viterbi) (في بعض الأحيان يمكن الحصول على نفس النتيجة بزيادة ٢٪ من الحالات الواقعة في فضاء البحث) فإن الأنظمة ذات العدد الكبير جدا من الكلمات قد تحتاج إلى طريق أكثر كفاءة. ونذكر على سبيل المثال الطرق التالية بدون الدخول في تفاصيلها.

يمكن استخدام البحث عن طريق العمق أولا (Depth first)، وهذه الطريقة يمكن أن تؤدي إلى كفاءة شديدة للبحث، ولكن يلزم مقارنة مسارات ذات أطوال مختلفة مما يُصَعِّب من التحكم في البحث. ويمكن استخدام طرق تعتمد علي المحولات محدودة الحالة «finite state transducers»، وهذه الطرق تعبر عن كل المعلومات المطلوبة للبحث، مثل قاموس النطق والنماذج الأكوستيكيَّة والنماذج اللغوية في صورة هذه المحولات، ثم يتم دمجها، ومن ثم تطبيق خوارزمات للتحديد (determinization)، وضغط (minimization) المحول الناتج للوصول إلى محول يمكن البحث فيه بكفاءة.

ومع أن محركات البحث مصممة في الأساس للحصول على أفضل جملة، فبالإمكان الحصول على عدد ميم (M) من الجمل مرتبة حسب احتماليتها بتعديلات بسيطة، ويمكن تخزين الجمل الناتجة بصورة أكثر كفاءة في ما يسمى «التشبيكة» (lattice).

ومن المتعارف عليه في البحث حاليا أن يتم استخدام محركات البحث متكررة-المرور (Multiple - pass). وفي هذه المحركات يتم البحث على عدة مراحل؛ ففي المرحلة الأولى يتم البحث باستخدام نماذج أكوستيكية ولغوية بسيطة نسبيًّا (على سبيل المثال ثنائي -جرام وثلاثي- فونات غير عابرة للكلمة) للحصول على (التشبيكة) ثم يتم إعادة تقييم (التشبيكة) باستخدام نماذج أكثر تعقيدا (مثال الرباعي - جرام وغيرها)؛ وتعتبر هذه الطريقة وسيلة معروفة للحصول على كفاءة عالية للبحث بدون التخلي عن الدقة. وبالإضافة إلى (التشبيكة) فإنه يمكن التعبير عن الحلول المتعددة باستخدام ما يسمى شبكة الاختلاط (confusion network)؛ وهذه يمكن اعتبارها أكثر كفاءة من (التشبيكة)، ولكن المسارات المتوازية فيها لا تعبر عن نفس الفترة الزمنية. وتستخدم

شبكات الاختلاط في تطبيقات متعددة؛ على سبيل المثال دمج نتائج عدد من نظم التعرف على الكلام. ويسمى ذلك طريقة الانتخاب.

وبعد عرض طرق التدريب والبحث لنظم التعرف على الكلام، فإننا سنعرض فيما يلي طريقة التدريب والبحث لبناء نظام التعرف على الكلام، ثم نتحدث ببعض الإسهاب عن بناء أنظمة التعرف على الكلام في اللُّغة العربية.

٣- ملخص لبناء نظام للتعرف على الكلام

لبناء نظام حديث للتعرف على الكلام فإنه يلزم الآتي:

٣,١ - البيانات الأكوستيكية (Acoustic data)

وتتكون من ملفات الكلام والنص المصاحب لها، وعادة تتكون هذه الملفات من عدة آلاف من الساعات للحصول على الدقة المطلوبة. ويمكن استخدام عدة مئات من الساعات في البداية ثم استخدام النظام الناتج للحصول على النص المناظر لبقية البيانات.

يتم الحصول على البيانات الأكوستيكية على النَّحو الوارد فيما سبق. ويتم بناء نماذج الفون، وعادة تستخدم نماذج ثلاثية - الفون إلي سباعية الفون العابرة للكلمات (Cross-word)، كما يتم استخدام شجرة القرار لربطها على النَّحو الَّذي سبق تفصيله. وفي المرحلة الأولى يتم الحصول على معاملات النماذج عن طريق تعظيم الاحتمالية (Maximize the probability) ثم تأتي الطرق التمييزية، وهي واسعة الانتشار في النظم عالية الدقة؛ ولكننا لن نتعرَّضَ لها في هذا الفصل، ويمكن الرجوع إلى بعض المراجع المدرَجة في نهاية الكتاب إذا أراد القارئ التعرف على هذه الطرق. وتعتبر حزمة البرامج HTK من أهم الأدوات في هذا الصدد.

٣,٢ - البيانات اللغوية

وتتكون هذه البيانات من النصوص. ويفضل أن تكون النصوص المستخدمة قريبة من الجُمل التي سيتم التعرف عليها لاحقا. فللتعرف على الأخبار يُفضل استخدام نصوص مستقاة من الأخبار - كالصحف أو المدونات. وتعتبر نصوص «جيجا وورد»

(Giga word) من أشهر هذه النصوص. وتستخدم البيانات اللغوية لبناء النماذج اللغوية كما تم تفصيله فيما سبق. وتعتبر حزمة برامج (SRI tool kit) من أهم الأدوات لبناء النماذج اللغوية.

٣,٣ - البحث

بعد بناء النماذج الأكوستيكية واللغوية يتم دمجها مع قاموس النطق لبناء فضاء البحث. وعند التعرف على الكلام فإنه يتم استكشاف فضاء البحث للوصول إلى أفضل سلسلة من الكلمات. وفي هذا الصدد – وكما ذكرنا سابقا – يستخدم عادة أحد محركات البحث متكررة المرور للوصول إلى أفضل الحلول بكفاءة. ونُشير إلى المحركات القائمة على المحولات محدودة الحالة، ويمكن استخدام حزمة البرامج من AT&T لهذا الغرض.

ولكن محركات البحث في العموم – وخاصة ذات الكفاءة العالية – لا تتواجد بشكل مفتوح (Open source)، ولكن يمكن استخدام محرك البحث الموجود مع HTK كبداية لدراسة الجانب العملي لتطوير محركات البحث ذات الكفاءة العالية.

وتكمن أهمية الطرق السابقة في أنها تعمل لأي لغة وبشكل آليّ، ولا تحتاج على الأقل نظريًّا لأي دراية باللغة، وربما يكون المكون الذي يحتاج إلى دراية باللغة في هذه النظم هو القاموس الصوتي.

وعلى الرغم من ذلك فإن الدراية بلغة ما واستخدامها بشكل أو بآخر يمكن أن يؤدي إلى تحسين نتائج التعرف على الكلام بشكل كبير؛ وفيما يلي سنتحدث باستفاضة عن التعرف على الكلام في اللُّغة العربية.

٤ - التعرف على الكلام في اللُّغة العربيَّة

نتحدث في هذا الجزء عن التعرف على الكلام في اللُّغة العربية. ومن حسن الحظ أن اللغة العربية كانت تحوز اهتمام مجتمع الباحثين في التعرف على الكلام خلال الفترة من ٢٠٠٤ إلى ٢٠١٠. وكان هذا متزامنا مع برامج (DARPA) لهذا الغرض. وتم بناء أنظمة للتعرف على الأخبار والمحادثات التلفونية وكذلك بعض اللهجات المحلية مثل العراقية والشامية. وأدى هذا الاهتمام لتطور نظم التعرف على الكلام العربية تطورًا كبيرًا.

وسنعرض - فيما يلي - لبعض خصائص اللغة العربية التي تمَّت معالجتها أثناء الأبحاث السابقة.

٤ , ١ - غياب التشكيل من النصوص العربية

من المعروف أن النصوص العربية الحديثة تُكتب بدون تشكيل. فعلى سبيل المثال فإن الكلمة «كتب» يمكن أن تكون «كَتَبَ» أو «كُتُب» أو غيرها. ويمكن للقارئ في أغلب الأحوال أن يستنتج النطق الصحيح من السياق. ورغم أن هذه الطريقة واسعة الانتشار فإنها تشكل تحديًا لنظم التعرف على الكلام لصعوبة أو ربما استحالة الحصول على نطق الكلمة دون وجود التشكيل.

ومع بداية الاهتمام بنظم التعرف على الكلام العربيّ فقد تم استخدام حروف الكتابة. ففي المثال السابق يكون نُطق كلمة «كتب» هو «كَتَبَ»؛ ومع أن هذه الطريقة التي تعتمد على الجرافيمات (Graphemes) (الجرافيم هو وحدة التَّحليل الكتابيّ/ الجرافيميّ) قد ساعدت على سرعة تطوير النظم، وخاصة بالنسبة لغير الناطقين بالعربية، فمن الواضح أن ذلك يتم على حساب الدقة. فنماذج الصوامت، مثل «ت» تمتص الحركات الصغيرة.

وكذلك فإن نماذج الصَّوائت، مثل «و» تخلط بين «و» الحر (الَّذي يُمَثِّلُ حرفَ مَدّ) مثل «نور» و «و» شبه المتحرك (الَّذي يُمَثِّلُ حرفَ لِين) مثل «دواء». ولبناء نماذج فونية للغة العربية فإنه يجب إضافة الحركات القصيرة. ويعتبر محلل (Backwalter) الصرفي من أهم الأدوات للقيام بذلك. فعند إدخال عدد من الكلمات إليه فإنه يعطي لكل كلمة (بالإضافة إلى مخرجات أخرى) كل طرق النطق المختلفة. وعلى الرغم من أنه لا يمكنه تحديد طريقة واحدة لكل كلمة في البيانات الأكوستيكية، فإنه يمكن بناء قاموس صوتي بطرق نطق متعددة. وباستخدام هذا القاموس يمكن بناء نماذج أكوستيكية فونية.

وفي مقارنة بين نماذج الجرافيم والنماذج الفونية للغة العربية الفصحى المعاصرة. (Modern Standard Arabic MSA) وُجِدَ أن النماذج الفونية تعطي «تفوقًا نسبيًّا» ١٠٪ (يجب التنويه أن التفوق النسبي ١٠٪ يعني أنه إذا كانت نسب الخطأ لنماذج الجرفيم هي ٢٠٪ فإن نسبة الخطأ لنماذج الفون تكون ١٨٪ وليس ١٠٪). ومن المفيد التنويه أيضا أن القواميس الصوتية كثيفة النطق - كما في حالة التشكيلات العربية (في

المتوسط محلل Backwalter) - تعطي حوالي ٦ تشكيلات مختلفة لكل كلمة)، وعندها يُفضل استخدام ما يسمى احتمالات النطق (Pronunciation Probabilities)، والتي يمكن استنتاجها ببساطة أثناء تدريب النظام.

ومن المفيد أن نلفت النظر إلى أن أي محلل صرفي قد يفشل في حل بعض الكلمات، خاصة الكلمات ذات الأصول الأجنبية أو التي لا توجد في قاعدة بياناته. وفي هذه الحالة يجب إيجاد تشكيل هذه الكلمات بطريقة ما سواء يدويا أو آليًّا؛ ومن المدهش أن بعض أكثر النظم نجاحا يقوم بمزج الفونات مع الجرافيمات للكلمات التي لا يوجد لها تشكيل.

وعلى الرغم من تفوق النماذج الفونية للعربية الفصحى فإن تطبيقها للتعرف على اللهجات المحلية، مثل: العراقية، لم يحالفه النجاح؛ ويرجع ذلك في الغالب إلى أن المحللات الصرفية - مثل Buckwalter - مصممة للتعامل مع اللغة الفصحى؛ وعادة ما تؤدي إلى تشكيل خاطئ للكلمات العامية، وخاصة الشائعة؛ وسنعود إلى هذه النقطة عند مناقشة التعرف على اللَّهجات العامية.

٤,٢ - البناء الصرفي للغة العربية

من المعروف أن اللغة العربية غنية صرفيًّا، مقارنة باللغات الأوروبيَّة كالإنجليزية. فبإضافة السوابق واللواحق يمكن تحويل أي كلمة إلى عدد كبير من الكلمات ذات الدلالات المختلفة؛ فكلمة «سيكتبه» - على سبيل المثال - تكافئ الجملة الإنجليزية «He will write it»؛ ولأن تعريف «الكلمة» في نظم التعرف هو سلسلة متصلة من الحروف، فإن استخدام السوابق واللواحق يؤدي إلى ظهور عدد كبير جدا من الكلمات المختلفة. وعلى سبيل المثال فإن قاموسًا مكونًا من ٦٤ ألف كلمة يكفي لتغطية ٩٩٪ من نصوص الأخبار في اللغة الإنجليزية، بينما نحتاج إلى عشرة أمثال هذا العدد أي حوالي ٦٠٠ ألف كلمة للوصول إلى نفس النسبة في اللغة العربية. وفي الحالة العامة لا يمثل ذلك مشكلة إذا توفرت الأدوات المناسبة لبناء فضاء البحث وكذلك محرك بحث ذو كفاءة عالية جدا للتعامل مع هذا العدد الضخم من الكلمات. ويمكن عن طريق التحليل الصرفي - أو حتى بعض الطرق البسيطة - فصل السوابق واللواحق أو حتى بعضها، ويؤدي هذا بطبيعة الحال إلى تقليص عدد الكلمات، وبالتالي تحسين التغطية للنصوص؛ ولكنه لا يؤدي بالضرورة إلى تحسين نسب التعرف على

الكلمات. ويمكن تفسير ذلك بأن فصل السوابق واللواحق يؤدي إلى تقليص القدرة التنبؤية للنماذج اللغوية، لأن السابق أو اللاحق في حالة فصله يعد ككلمة منفصلة، في حين أن قدرته التنبؤية تكون ضعيفة للغاية (فالسابق «ال» – على سبيل المثال – يمكن أن يأتي بعده عدد كبير جدا من الأسماء في اللغة العربية)؛ كما أن بقاء الكلمة كوحدة متصلة يؤدي إلى نماذج أكوستيكية أكثر استقرار يسهل تمييزها عن بقية الكلمات، ولكن هذا لا ينفي أن التحليل الصرفي قد يكون مفيدا في بعض الحالات مثل عدم وجود أدوات للتعامل مع فضاء بحث كبير جدا أو الرغبة في بناء قاموس صغير نسبيا أو حتى عدم وجود نصوص كافية للحصول على عدد كبير جدا من الكلمات. مما سبق يتضح أن اختيار القاموس يجب أن يحتوي على بُعد جديد، وهو التحليل الصرفي. وفي هذا الإطار يجب الإجابة عن أسئلة مثل: أي الكلمات ستخضع للتحليل الصرفي؟ وما هي السوابق واللواحق التي سيتم اختيارها؟ وكيف يمكن بناء النماذج اللغوية في هذه الحالة؟؛ ونعتقد أن الإجابة عن الأسئلة السابقة تعتمد على كمية البيانات المتاحة، وكذلك طبيعة النظام المستهدف؛ ويجب إجراء تجارب للوصول إلى أحسن تصميم للقاموس.

٤, ٣ – التعرف على اللهجات العامية

بالإضافة إلى اللغة الفصحى المعاصرة التي تستخدم في الكتب والصحف ووسائل الإعلام، فإن للغة العربية لهجات عامية (دارِجة) مختلفة، مثل: الشامية والخليجية. ومع أن معظم اللغات لها صبغة رسمية وأخرى عامية، فإن الاختلاف في اللغة العربية يمكن اعتباره كبيرًا جدًّا.

ومع البُعد عن المناقشة الفلسفية في الفرق بين اللهجة واللغة، فإننا نجد الاختلاف بين اللهجات العربية يتخطَّى الكلمات إلى الفونات؛ فعلى سبيل المثال، لا توجد «الجيم القاهرية» في كثير من اللهجات الأخرى. وكذلك فإن الكثير من الكلمات الشائعة في لهجة ما لا توجد في لهجات أخرى .

ومما يزيد المشكلة تعقيدا أن اللهجات لا تكون مكتوبة، ولذلك لا توجد نصوص كافية لبناء النماذج اللغوية. وقد تم بناء نظم للتعرف علي اللهجات العربية المختلفة مثل المصرية والعراقية والشامية بنسبة خطأ ٣٠٪ ومما لا شك فيه أنَّه مازال يلزم الكثير من العمل للحصول علي نظم تعرف علي اللهجات ذات دقة كبيرة.

المبحث الثَّاني
نُظُم تحويل النَّصّ إلى كلام

١- التوصيف اللغوي.

٢- إنتاج إشارات الكلام.

نُظُم تحويل النَّصّ إلى كلام
(Text to Speech - TTS)

إن نظم تحويل النص إلى كلام - كما يوحي اسمها - تقبل النص في لغة ما وتنتج ذبذبات الكلام المناظرة لهذا النص. وفي أبسط صورها يمكن لهذه النظم تسجيل الإشارات المناظرة للنص المراد نطقه؛ ولكن هذه الطريقة تظل مرتبطة بنطق عدد محدود من الجمل.

لذلك فإن الصورة العامة لنظم تحويل النص إلى كلام تعتمد على تخليق الكلام من بعض الوحدات الصوتية؛ ولذلك فإن الحصول على هذه الإشارات التخليقية بجودة تقترب من الكلام الطبيعي هو الهدف الأساسي لهذه النظم.

وعادة تنقسم نظم تحويل النص إلى كلام إلى جزأين أساسيين، يقوم الجزء الأول بتحويل النص إلى «توصيف لغوي». ويقوم الجزء الآخَر بتحويل التوصيف اللغوي إلي إشارات الكلام؛ ويكون الجزء الأول معتمدا على اللغة ويحتاج إلى الكثير من الخبرة اللغوية في حين أن الجزء الآخَر لا يعتمد على اللغة، وإنَّما يعتمد في الأساس على البيانات المسجلة (بالطبع فإن البيانات المسجلة يجب أن تكون من اللغة المراد نطقها).

وسنقوم فيما يلي بعرض سريع لتكوين التوصيف اللغوي دون الدخول في تفاصيل مع التركيز على الجزء الآخر، وهو المعنِيّ بإنتاج إشارات الكلام من التوصيف اللغوي. ولإنتاج الكلام تاريخ طويل يعود إلى نظم الـ (Formant) والدايفونات.

أمَّا في هذا الفصل، فسوفَ نتناول الطرق الكثيفة الاستخدام للبيانات؛ وهي طرق اختيار الوحدات، والطرق الإحصائية التي تستخدم نماذج ماركوف المخفِيَّة، وهي واسعة الانتشار الآن.

١ - التوصيف اللغوي

يتكون التوصيف اللغوي - في أبسط صورة - من تحويل النص إلى سلسلة من الفونيمات. أي: يجب أن تتوافر أداة لتحويل النص إلى فونيمات مما يسمى عادة بأدوات الجرافيم - إلى- فونيم (Grapheme to phoneme) وهذه الأدوات تكون إما لغوية باستخدام قاموس وبعض القواعد، أو إحصائية قائمة على التدريب باستخدام الأمثلة؛

وفي بعض الأحيان يتم المزج بين القواعد اللغوية والإحصائية؛ وفي اللغة العربية - حيث يُكتب النص في أغلب الأحيان بدون التشكيل - فإن وجود أداة لإضافة التشكيل قبل التحويل لفونيمات تعد أساسية للحصول على التوصيف اللغوي. وعادة يتم معالجه النص قبل تطبيق أدوات الجرافيم - إلى - فونيم. فعلى سبيل المثال يتم تحويل الأرقام والتواريخ وعلامات الترقيم إلى نص؛ وبالإضافة إلى تحويل الجرافيمات إلى الفونيمات فإن التوصيف اللغوي يحتوي على معلومات عديدة من نفس الكلمة، مثل: الفونيمات المجاورة وعدد المقاطع في الكلمة وموقع المقطع، وكذلك معلومات عن الجملة، مثل: موقع الكلمة في الجملة وعلامات الترقيم المستخدمة. وبالإضافة إلى ذلك فإنه يتم استخدام معلومات عن المتحدث ونوع الأسلوب (خبري أو استفهامي) وحالة المتحدث.

وتتكامل هذه المعلومات مع سلسلة الفونيمات لاختيار أكثر الوحدات مناسبة لتخليق الكلام. فعلى سبيل المثال، تُستخدم هذه المعلومات لاستنتاج امتداد الوحدة والتردد الأساسي لها. ومما سبق يتضح أن أدوات التوصيف اللغوي يمكن تقسيمها إلى ما يلي: أدوات لمعالجة النص، مثل: التشكيل الآلي، وأدوات لتحويل الجرافيم إلى فونيم، وأدوات للتعامل مع المعلومات على مستوى الكلمة والجملة والمتحدث. وترتبط هذه الأدوات ارتباطا وثيقا باللغة التي يتم التعامل معها، وتتطلب معرفة دقيقة بخصائصها وكيفية توصيفها.

-ولخصوصية هذه القواعد واختلافها من لغة إلى أخرى فإننا لن نتطرق إليها بالتفصيل.

٢- إنتاج إشارات الكلام

كما أشرنا سابقا، فإن إنتاج إشارات الكلام الحديثة يتمُّ باستخدام إحدى طَرِيقَتَين؛ طريقة اختيار الوحدات والطريقة الإحصائية؛ وسنتعرض لكلتا الطريقتين فيما يلي:

٢,١- طريقة اختيار الوحدات

تعتمد طريقة اختيار الوحدات - كما يوحي الاسم - على تسجيل قاعدة بيانات من متكلم واحد وتقسيمها إلى وحدات؛ ويتم إنتاج الكلام باختيار الوحدات المناسبة. وسنفصل فيما يلي كيفية الحصول على الوحدات وكيفية اختيارها.

يتم تسجيل عدة ساعات، أي عدة آلاف من الجمل من متكلم واحد، ويُراعى الحصول على تسجيل نقي، كما يُراعى اختيار الجمل بحيث تناسب ما سيتم نطقه فيما بعد، ثم تُقَسَّم الجمل إلى وحدات فونيمية؛ وعلى الرغم من أن وسيلة التقسيم ليست ذات أهمية كبرى، فإنَّ نماذج ماركوف المخفِيَّة تستخدم في هذا السياق. ونُشير إلى أن استخدام نماذج ماركوف هنا فقط لتقسيم الكلام وليست لإنتاجه كما في الطريقة الإحصائية التي سيتم تفصيلها لاحقا.

يتم تسجيل عدة آلاف من الجمل من متحدث واحد، ويتم بناء النماذج الفونيمية. في حقيقة الأمر يتم بناء نماذج تعتمد على السياق، مثل التراي فون أو غيره كما أورَدنا في المبحث المعنيّ بالتعرف على الكلام.

وللتذكرة فإن ناتج هذه النماذج يكون شجرة القرار لكل فونيم، والتي تعبر عن هذا الفونيم في السياقات المختلفة، حيث تمثل كل ورقة فيها وحدات هذا الفونيم التي تتشارك في السياق - أو بعبارةٍ أخرى - التي لها سياقات متشابهة؛ ويتم تمثيل كل ورقة بخليط (جاوس) كما سبق أن فصلنا.

وبعد بناء هذه النماذج، يمكن استخدامها لتقسيم إشارات الكلام إلى فونيمات، أي معرفة متى يبدأ وينتهي كل فونيم في الإشارة المعطاة. وتُعرَفُ هذه العملية بـ «التقسيم» (segmentation). وبعد المرور على إشارات الكلام كلها يتم ربط عدد من المقاطع بكل ورقة من أوراق شجرة القرار لكل فونيم. فعلى سبيل المثال، بعد إجراء التقسيم يتم تجميع كل المقاطع التي تنتمي إلى الفونيم / ب/ . ولكل مقطع يتم المرور على شجرة القرار المناظرة للوصول إلى الورقة المناسبة. وللتذكرة فإن الوصول إلى الورقة يتم عن طريق الإجابة عن أسئلة خاصة بالسياق (الفونيمات المجاورة في أغلب الأحيان). وعلى سبيل المثال، إذا اعتبرنا الفونيم / ب/ في كلمة «كبير» فإنه يمكن الوصول إلى الورقة المناسبة عن طريقة الإجابة عن أسئلة السياق الفونيمي (الفتحة / و / ى/ في هذه الحالة)

اختيار المقاطع (الوحدات) لإنتاج الكلام

سنفصل فيما يلي كيفية إنتاج كلمة «كبير» كما في المثال السابق. وسنفترض أن لكل فونيم شجره قرار واحدة، وليس لكل حالة من الفونيم للتسهيل، ولأن ذلك أكثر شيوعا في نظم إنتاج الكلام.

إن كلمة «كبير» تتكون من السلسلة الآتية من الفونيمات:

/k/ /a/ /b/ /I/ /r /

ويمكن كتابتها كتريفونات كما يلي:

/#-k-a/ /k-a-b/ /a-b-I/ /b-I-r/ /I-r-#/

حيث # هو فونيم يعبر عن بداية ونهاية الكلمة، وحيث يؤخَذ سياق الفونيم في الاعتبار عند كتابة كل تريفون. ويتم إضافة الفترة (duration)، والنغمة (pitch) لكل تريفون. ويمكن حساب الفترة المستهدفة والنغمة المستهدفة باستخدام عوامل كثيرة اعتمادًا على التوصيف اللغوي المصاحب وموقع الفونيم في الكلمة والموقع في الجملة وعلامات الترقيم وحالة المتحدث. ولن نتطرق بالتفصيل لهذه النقطة لعدم وجود دراسة قياسية شائعة الانتشار، بالإضافة إلى أن ذلك يعتمد بدرجة كبيرة على اللغة وعلى التوصيف اللغوي؛ كما أننا نعتقد أن اللغة العربية بحاجة لكثير من العمل في هذا المجال. وبعد إضافة الفترة والنغمة المستهدفة لكل تريفون يمكن المرور علي شجرة القرار المناسبة والوصول إلى الوحدات المرتبطة بالورقة المناسبة .

ويتم اختيار الوحدات المناسبة لإنتاج الكلام عن طريق تقليص التكلفة. ويتم حساب التكلفة كالآتي:

التكلفة الكلية = تكلفة الفترة + تكلفة النغمة + تكلفة اللصق

وحيث نحسب تكلفة الفترة والنغمة حسب بعدهما عن الفترة والنغمة المستهدفين؛ ونحسب تكلفة اللصق بحيث تتيح انتقالا سلسًا للكلام.

٢ , ٢- الطريقة الإحصائية: نماذج ماركوف المخفِيَّة

لقد عرضنا فيما سبق نماذج ماركوف المخفِيَّة في مجال التعرف على الكلام؛ ولكن يمكن استخدامها أيضا في إنتاج الكلام. فعلى سبيل المثال نعتبر التريفون / k–a–b/ ثلاثي الحالة وكذلك نماذج جاوس البسيطة. إن استخدام هذا النموذج لإنتاج الكلام يؤدي في أبسط صورة لاستخدام متجهات المتوسط الحسابي لكل حالة مكررة حسب احتمالات الانتقال لكل حالة. ولكن هذه الطريقة المبسطة تؤدي إلى جودة متدنية للكلام للسَّبَبَين التَّاليين:

- تكرار المتوسط الحسابي بدون أخذ السياق في الاعتبار.
- تكلفة عالية للصق لعدم اعتبار الوحدات المجاورة.

وللتغلب على هذا فقد تم استخدام احتمالات المشتقة الأولى والثانية في إنتاج الكلام. ومعاملات المشتقة واسعة الانتشار في التعرف على الكلام هي عبارة عن مزج خطي بين متجهات الطيف. وعلى سبيل المثال فإن المشتقة الأولى عند الزمن (t) للكبسترم C(t) يمكن كتابتها كما يلي:

$$DC(t) = \sum_{k=-2}^{+2} kC(t-k)$$

وكما يتضح فإن هذه المعاملات تأخذ السياق والوحدات المتجاورة ضمنيا في الاعتبار. ودون الدخول في التفاصيل الحسابية فإن إنتاج الكلام باعتبار المشتقات يؤول إلى حل مجموعة معادلات من الدرجة الأولى؛ وبما أن تدريب نماذج ماركوف يتم عادة في فضاء الكبسترم أو أي فضاء آخر مشابه، ولا يتم في فضاء الإشارة الزمنية، فإنه يلزم وجود مرشِّح لإنتاج إشارات الكلام بدءً من المتجهات المنتجة. ويمكن استخدام المرشحات المستخدمة في تحليل الكلام لهذا الغرض.

فعلى سبيل المثال، بمعرفة معاملات التنبؤ الخطي (linear prediction Coefficients)، والاستثارة (excitation) يمكن إنتاج الكلام باستخدام مرشح التنبؤ الخطي المعروف. وفي أول أعمال إنتاج الكلام باستخدام النماذج الإحصائية تم استخدام مرشح قائم على معاملات (كبسترم ميل) (MFCC) المعروفة بجودتها لبناء نماذج (ماركوف). وكما ذكرنا في سياق هذا الفصل، فإن إنتاج الكلام يحتاج إلى نماذج للنغمة والفترة. وبما أن للنغمة طبيعة منفصلة، فإنه عند المزج بين معاملات الطيف والنغمة يتم استخدام نماذج تمزج بين التوزيعات المتصلة والمنفصلة. إن نماذج ماركوف تستخدم ضمنيا نماذج هندسية للفترة؛ ومن المعروف أنها غير دقيقة في التنبؤ بفترة الوحدات؛ وبينما لا يمثل ذلك مشكلة كبيرة للتعرف على الكلام فإنه من الأفضل استخدام نماذج أكثر دقة عند إنتاج الكلام. لذلك يتم - في بعض الأحيان - استخدام ما يسمى بنماذج شبه- ماركوف المخفِيَّة (Hidden semi-Markov Models).

ولتلخيص ذلك فإنه عند إنتاج الكلام يتم عادة بناء نماذج شبه- ماركوف المخفِيَّة والمزج بين التوزيعات المنفصلة لمعاملات النغمة المتصلة للطيف، ويتم استخدام المشتقات لفرض الاتصال حين إنتاج إشارات الكلام.

وفيما يلي سنبين كيفية إنتاج إشارات الكلام لكلمة «كبير». وكما سبق فإننا نقوم بكتابة التراي فونات المناظرة.

/#-k-a/ /k-a-b/ /a-b-I/ /b-I-r/ /I-r-#/

ولنأخذ التراي فون /a–b–I/ كمثال. لكل حالة من التراي فون يتم التنبؤ بعدد متجهات السِّمات حسب توزيع الفترة للحالة، ويتم إنتاج متجهات السِّمات لكل من متجهات الطيف والنغمة باستخدام توزيعات التراي فون مع أخذ المشتقة الأولى والثانية في الاعتبار باستخدام المتجهات والنغمة والمرشح المناسب، وعليه يتم إنتاج إشارات الكلام، ويتم تكرار ذلك لكل تراي فون.

وكما ذكرنا سابقا، فإن طريقة اختيار الوحدات تؤدي إلى كلام عالي الجودة إذا توفرت قاعدة بيانات كافية لتغطية السياقات المختلفة، ولكنها تفتقد المرونة في تغيير نوع الكلام أو المتكلم.

وتقع الطريقة الإحصائية على الجانب الآخر، فهي تتيح مرونة كبيرة لتغيير المتكلم أو نوع الكلام باستخدام التحويلات الخطية واسعة الانتشار في نظم التعرف على الكلام، ولكنها في نفس الوقت - ونتيجة لأخذ المتوسطات أثناء التدريب- قد تؤدي إلى كلام أقل جودة من أفضل نظم اختيار الوحدات؛ ولكنها تبقى طريقة ذات دقة مناسبة ومرنة في نفس الوقت.

ولتحسين جودة الطريقة الإحصائية فقد تم دراسة بعض الطرق في السنوات الأخيرة، من أهمها:

- طريقة مصفوفة التباين العامة (Global Covariance Matrix)، وذلك لتلافي تأثير المتوسطات في إنتاج الكلام.
- تدريب النماذج باستخدام أقل خطأ في التوليد.

المبحث الثَّالث
نُظُم التَّعَرُّف على اللُّغة والمتَكَلِّم

١- نظم التعرف على اللُّغة.

٢- نظم التعرف على المتكلم.

في هذا المبحث سنعرض نظم التعرف على اللغة ثم نشرح نظم التعرف على المتكلم. ولهذه النظم تطبيقات عديدة في مختلف المجالات.

١- نظم التعرف على اللُّغة

إن الهدف في نظم التعرف على اللغة هو معرفة أي لغة نُطِقت بمعرفة ذبذبات كلامٍ ما. وتكون المدخلات للنظام هي عدة ثواني من ذبذبات الكلام، يحدد النظام على أساسها اللغة التي نطقت به من بين عدة لغات معروفة لديه. وفي بعض الأحيان تشمل هذه النظم نظم التحقق من اللغة، بمعنى أن لدينا إشارة كلام ولغة ما ونريد التحقيق ممَّا إذا كانت هذه الذبذبات تنتمي لهذه اللغة أم لا. ولكن في هذا الفصل سنكتفي بعرض نظم التعرف على اللغة لتشابه الطرق المستخدمة.

باعتبار ذبذبات الكلام - أو لنكون أكثر دقة متجهات السِّمات المناظرة - X فإنه يمكن صياغة مسألة التعرف على اللغة كما يلي:

$$l^* = \underset{l}{argmax}\ P(l|x)$$

حيث L مجموعة من اللغات المعرَّفة للنظام. وهكذا تصبح المسألة تطبيقًا بسيطًا لقواعد التعرف على الأنماط.

ومن النماذج الشائعة في هذا المجال استخدام نماذج خليط (جاوس) لكل لغة من اللغات المراد التعرف عليها، ويتم تدريب هذا الخليط من متجهات السِّمات التي تنتمي لكل لغة.

ويمكن استخدام هذه النماذج لحساب الاحتمالات في المعادلة السابقة عند التعرف على اللغة، ويسمى هذا الأسلوب «الأسلوب الأكوستيكي» لأنه يستخدم الإشارات الأكوستيكية بشكل مباشر. ومن المعروف أن هذا الأسلوب لا يعطي نتائج جيدة جدا للتعرف على اللغة، وربما يعود ذلك لتداخل عوامل كثيرة - بالإضافة إلى اللغة - في الإشارة الأكوستيكية، مثل: المتكلم والتسجيل. وبجانب الأسلوب الأكوستيكي، فإن ما يسمى بـ «الأسلوب الفونوتيكي» واسع الانتشار في نظم التعرف على اللغة.

وفي الأسلوب الأخير يتم استخدام نظام للتعرف على الفونيمات. ومن المدهش أن هذا النظام لا ينتمي بالضرورة إلى أي من اللغات المراد التعرف عليها. وبعد تمرير الإشارات في اللغات المراد التعرف عليها على نظام التعرف الفونيمي، فإنه يتم بناء نماذج النَّحو الإحصائي (N-gram) لكل لغة على الفونيمات الناتجة. والفكرة الأساسية هنا أن سلاسل الفونيمات الناتجة تستطيع التمييز بين اللغات المراد تمييزها.

ومن البديهي أنه يمكن استخدام هذه الطريقة للتمييز بين أي عدد من اللغات. ومن المعروف أن الطريقة الفونيتيكية تؤدي إلي نتائج ممتازة في التعرف على اللغة، ربما لأنها أقل تأثرا من الطريقة الأكوستيكية بالمتغيرات، مثل المتحدث والتسجيل.

ويمكن تعميم وتحسين الطريقة الفونوتيكية باستخدام عدة نظم للتعرف الفونيمي، وليس بالضرورة من اللغات التي يراد تمييزها، ثم بناء نماذج (N-gram) اللغوية لكل لغة ونظام فونيمي.

فعلى سبيل المثال، للتمييز بين اللغة «أ» واللغة «ب» يمكن استخدام نظام فونيمي للغة «ج» واللغة «د»؛ ومن ثم بناء نماذج لغوية «أ جـ» و «أ د» و «ب جـ» «ب د». ويسمى هذا النظام «النظام الفونوتيكي المتوازي»، ويؤدي إلى نتائج أفضل من النظام الفونوتيكي، ولكن بالطبع مع زيادة حجم النماذج والعمليات الحسابية المصاحبة.

وعلى الرغم من أن الأسلوب الفونونيكي يبدو الأفضل في نظم التعرف على اللغة، فإن مزج الأسلوبين الفونومنيكي والأكوستيكي يؤدي في أغلب الأحيان إلى نتائج أفضل، ويتم هذا المزج بطرق عديدة. ففي أبسط صورة يمكن بناء نظامين منفصلين (أحدهما فونونيكي والآخر أكوستيكي) ومزج نتيجة النظامين للحصول على النتيجة النهائية.

٢- نظم التعرف على المتكلم

تَستَخدِمُ نظمُ التعرف على المتكلم - كما يوحي الاسم- إشارات الكلام للتعرف على المتكلم. وتنقسم هذه النظم إلى نوعين؛ في النوع الأول يكون لدينا مجموعة من المتكلمين ويهدف النظام إلى معرفة إلى أي منهم تنتمي إشارة الكلام المدخلة. أما في النوع الآخَر، وهو التحقيق، فإنه يعطي لنا متكلمًا ويهدف النظام إلى معرفة إذا كانت

إشارة الكلام المدخلة صدرت من هذا المتكلم أم لا. وللنوع الثاني تطبيقات كثيرة إذ يمكن استخدامه للتحقق من البيانات مثل استخدام بصمات الأصابع أو العين وكذلك يمكن استخدامه بجانب كلمة السر (password) وتتشابه الطرق والنماذج المستخدمة في كلا النوعين ولذلك أننا سنتطرق إلى الطرق المستخدمة في التحقق من المتكلم.

وللتحقق من المتكلم S باستخدام إشارة الكلام X فإن الهدف يكون التأكد بأن X تنتمي إلى S أم لا. ولهذا يلزم وجود نموذج للمتكلم لحساب الاحتمال $P(X|S)$ وكذلك نموذج للخلفية (background) لحساب $P(X|background)$ أو للتسهيل $P(X|b)$ وعادة يتم الحكم بأن الإشارة X تنتمي إلى S إذا كان ناتج القسمة $\frac{P(X|S)}{P(X|b)}$ أكبر من قيمة محددة t تسمى العتبة (threshold).

ولفترة طويلة ظلت نماذج خليط (جاوس) تستخدم للتعبير عن المتكلم والخلفية. فيمكن تجميع عدد من متجهات السِّمات من المتكلم واستخدامها لبناء نموذج خليط (جاوس) بالطرق التقليدية.

وكذلك يمكن تجميع متجهات السِّمات من عدد كبير من المتكلمين واستخدامها لبناء نموذج الخلفية. ولتقليل كمية الكلام المطلوب الحصول عليه من المتكلم لتسجيله في النظام فإنه يمكن بناء نموذج خلفية من عدد كبير من المتكلمين ثم باستخدام طرق التكيف (adaptation) للحصول على نماذج المتكلم. ومن المدهش أن نفس متجهات السِّمات المستخدمة في التعرف على الكلام يتم استخدامها في نظم التعرف أو التحقق من المتكلم وهي معاملات الكبسترم ومشتقاتها. وعلى الرغم من أن نظم التعرف على الكلام تهدف إلى تحييد تأثير المتكلم فإن المتجهات المستخدمة فيها تؤتي نتائج طيبة جدا عند استخدامها في التعرف على المتكلم وتكون أفضل من الخصائص المعروفة بارتباطها ارتباطا وثيقا بالمتكلم مثل النغمة. ويمكن من الناحية اللغوية أو الأكوستيكية دراسة سِمات تؤدي إلى تحسين نظم التعرف أو التحقق من المتكلم.

وفي الجيل الأحدث من نظم التحقق من المتكلم يمكن النظر إلى المسألة على أنها مسألة تصنيف (classification)، ويمكن بناء مصنف لتحديد: هل تنتمي إشارات الكلام إلى متكلم معين أم لا.

وحديثًا تم استخدام آليَّات المتَّجهات الدَّاعمة (Support Vector Machine) لهذا الغرض. ومن الممكن بناء المصنف باستخدام متجهات السِّمات التي تماثل تلك المستخدمة في نماذج جاوس ولكن يمكن أيضا المزج بين نماذج خليط (جاوس) والمصنفات. فيمكن بناء نماذج خليط (جاوس) كما سبق أن ناقشنا ثم يلي ذلك بناء مصنفات في فضاء مكون من التوزيعات الاحتمالية لكل مكون من مكونات الخليط. وعادة تعطى الطريقة الأخيرة نتائج جيدة جدا للتحقق من المتكلم.

ولتقييم طرق التحقق من المتكلم فإنه يوجد نوعان من الأخطاء؛ في النوع الأول يتم قبول إشارة ما في حين أنها لا تنتمي إلى المتكلم المراد، ويسمى هذا النوع من الأخطاء: أخطاء التنبيه الخاطئ (false alarm)؛ وفي النوع الآخَر لا يتم قبول الإشارة بالرغم من أنها تنتمي إلى المتكلم المراد، وتسمي هذه بأخطاء الإخفاق (misrecognition error).

ويتم تصميم النظام عند نقطة تناسب التطبيق المستخدم لأصله. فمثلا عند استخدام النظام لوصول إلى الحساب البنكي فإن تكلفة التنبيه الخاطئ تكون أكبر كثيرا من تكلفة الإخفاق.

ببليوجرافيا مرجعيَّة

المبحث الأوَّل

1. AbuZeina, D.; Elshafei, M. (2011). Cross-Word Modeling for Arabic Speech Recognition. Springer.
2. Afify, M.; Nguyen, L.; Xiang, B.; Abdou, S.; Makhoul, J. "Recent progress in Arabic broadcast news transcription at BBN", in Proceedings of Interspeech, Lisbon, Portugal, September 2005.
3. Ali, A.; Zhang, Y.; Cardinal, P.; Dahak, N.; Vogel, S.; Glass, J. "A complete Kaldi recipe for building Arabic speech recognition systems", Proc. icassp 2014.
4. Bailly, G.; Perrier, P.; Vatikiotis-Bateson, E. (2012). Audiovisual Speech Processing. Cambridge University Press.
5. Chen S. F.; Goodman, J. (1999). "An empirical study of smoothing techniques for language modeling", Computer Speech and Language, vol. 13, pp. 359-394.
6. Clark, A.; Fox, C.; Lappin, S. (2010). The Handbook of Com-putational Linguistics and Natural Language Processing. John Wiley & Sons.
7. Cohen, I & Benesty, J.; Gannot, S. (2010). Speech Processing in Modern Communication: Challenges and Perspectives. Springer.
8. Dahl, G.; Yu, D.; Li, D.; Acero, A. "Context-dependent pre-trained deep neural networks for large vocabulary speech recognition", IEEE Transactions on Audio, speech and Language Processing", vol. 20, no. 1 , January 2012.
9. Elmahdy, M.; Gruhn, R.; Minker, W. (2012). Novel Techniques for Dialectal Arabic Speech Recognition. Springer.
10. Eskenazi, M.; Levow, G.; Meng, H.; Parent, G.; Suendermann, D. (2013). Crowdsourcing for Speech Processing: Applications to Data Collection, Transcription and Assessment. Wiley.

11. Gales, M.; Young, S. (2007). “The application of hidden Markov mod-els in speech recognition”, Foundations and Trends in Signal Pro-cessing, vol. 1, no. 3, pp. 195-304.

12. Mariani, J. (2010). Language and Speech Processing. John Wiley & Sons.

13. Mehler, A.; Romary, L. (2012). Handbook of Technical Communi-cation. Walter de Gruyter.

14. Mohammed, A. K. (2012). Speech Recognition System: Speaker De-pendent Recognizer for Sidama Language. Lambert Academic Publishing.

15. Mohri, M.; Pereira, F.; Riley, M. (2002). “Weighted finite state transducers in speech recognition”, Computer Speech and Lan-guage, vol. 16, no. 1, pp. 69-88.

16. Ortmanns, S.; Ney, H.; Aubert, X. (1997). “A word graph algorithm for large vocabulary continuous speech recognition”, Computer Speech and Language, vol. 11, no. 1, pp. 43-72.

17. Pathak, M. A. (2013). Privacy-Preserving Machine Learning for Speech Processing. Springer.

18. POvey, D.; Valtchev, V.; Woodland, P. C. (2006). The HTK Book (for HTK Version 3.4). University of Cambridge, http://htk.eng.cam.ac.uk, December 2006.

19. Rabiner, L. R. (1989). “A tutorial on hidden Markov models and selected applications in speech recognition”, Proceedings of IEEE, vol. 77, no. 2, pp. 257-286.

20. Russell, J.; Cohn, R. (2012). Speech Processing. Book on Demand.

21. Selouani, S. (2011). Speech Processing and Soft Computing. Spring-er.

22. Suh, J. (2012). Effective Data Selection Technology for Robust Speaker Recognition. University of Texas at Dallas. Graduate Pro-gram in Electrical Engineering.

23. Vasquez, D.; Gruhn, R.; Minker, W. (2013). Hierarchical Neural Network Structures for Phoneme Recognition. Springer.
24. Wang, J.; Furui, S.; Juang, B. (2012). Real World Speech Processing. Springer.
25. Sak, H.; Senior, A.; Beaufays, F. (2014). "Long Short-Term Memory Based Recurrent Neural Network for Large Vocabulary Speech Recognition" arXiv preprint, arXiv: 1402.1128, 2014.
26. Sundermeyer, M.; Schluter, R.; Ney, H. (2012). "LSTM Neural Network Language Modelling" InterSpeech 2012.

المبحث الثَّاني

1. Black, A. W.; Zen, H.; Tokuda, K. (2007). "Statistical parametric speech synthesis", Proc. of ICASSP, pp.1229-1232, Apr.
2. Gupta, S.; Bhatia, P. (2012). Text to Speech System: An Aid to Visually and Vocally Impaired. LAP Lambert Academic Publishing.
3. Hunt, A.; Black, A. (1996). "Unit selection in a concatenative speech synthesis system using a large speech database", Proceedings of ICASSP 96, vol. 1, pp. 373-376, Atlanta, Georgia.
4. Rao, K. S. (2012). Predicting Prosody from Text for Text-to-Speech Synthesis. Springer.
5. Rao, K. S.; Koolagudi, S. G. (2012). Emotion Recognition using Speech Features. Springer.
6. Van den Oord, A.; Dieleman, S.; Zen, H. et al (2016). "Wavenet: A Generative Model of Raw Audio", arXiv:1690.03499, Septmber 2016.

المبحث الثَّالث

1. Beigi, H. (2009). Fundamentals of Speaker Recognition. Springer.
2. Ghosh, D.; Debnath, D. (2012). Speaker Recognition: Real Time Veri-fication of VLSI Architecture Based on MEL Frequency Cepstral Co-efficients. Lambert Academic Publishing.
3. Hardcastle, W. J.; Laver, J.; Gibbon, F. E. (2012). The Handbook of Phonetic Sciences. John Wiley & Sons.
4. Heigold G.; Moreno, I.; Bengio, S. (2016). “End-to-End Text-Dependent Speaker Verification” ICASSP 2016.
5. Neamat El, G.; Yee, S. (2018). Computational Linguistics, Speech and Image Processing for Arabic Language. World Scientific.
6. Neustein, A. (2012). Forensic Speaker Recognition: Law Enforcement and Counter-Terrorism. Springer.
7. Patil, H. A. (2012). Advances in Speaker Recognition. Springer.
8. Rao, R. R.; Prasad, V. K. (2012). Automatic Text Independent Speaker Recognition Using Source Feature: Modeling Automation. Lambert Academic Publishing.
9. Reynolds, D. A.; Campbell, W. (2007). Text-Independent Speaker Recognition, Springer Handbook of Speech Processing and Communication, Springer-Verlag GMBH, Heidelberg, Germany.
10. Schultz, T.; Kirchhoff, K. (2006). Multilingual Speech Processing. Academic Press.

الفصل الثَّالث

تطبيقات مُعالجة اللُّغة العربيَّة في مجال التَّعلِيم

د. شريف مهدي عَبده

١ – تقنيات معالجة اللغات الطبيعية.

٢ – تعلم النطق باستخدام تقنية التعرف الآليّ على الكلام.

٣– تعلم الكتابة باستخدام تقنية التعرف الآليّ على الكتابة.

٤ – مقترحات بَحثِيَّة.

يُستخدم الحاسوب في تعلّم اللغات بصورة خاصة؛ لتعلّم مهارات اللغة، سواء أكانت اللغة الأم، أم اللغة الأجنبية، أو ما يسمى باللغة الثانية. وتُستخدم تكنولوجيا الحاسب الآلي أداة تعليمية تساعد متعلمي اللغة لتطوير مهاراتهم اللغوية، وتمثّل بذلك عنصرًا مكملاً بالإضافة إلى طرق تعليمية أُخرى، مما يساعد على خلق بيئة تعليمية نشطة وغنية لغويًّا. وتُعرّف موسوعة ويكيبيديا استخدام الحاسوب في تعلّم اللغات بأنّه استخدام تكنولوجيا الحاسب الآلي لتقديم وتعزيز وتقييم المادة المراد أنْ يتعلمها الطالب، وذلك من خلال الاستفادة من ميزات الحاسوب التفاعلية وأنماطه التعليمية والتعلّمية المختلفة والإنترنت.

وقد بدأ استخدام الحاسوب فعليًّا في تعلّم اللغات في الستينيَّات من القرن العشرين [٣٥]. وتطورت برامج تعلّم اللغة الإنجليزية بمساعدة الحاسوب مع بداية الثمانينيَّات، ومرَّ استخدام الحاسوب باعتبارِه مساعدًا في تعليم اللغات وتعلّمها بمراحل ثلاث، إذْ بدأت المرحلة الأولى فكرةً في الخمسينيَّات، وطُبِّقت في الستينيَّات، وقامت على أساس النظرية السلوكية التي عدّت الحاسوب أداة مثالية للتعليم، لأنَّه يسمح بتكرار تعلّم المادة مرات عديدة [٣٦].

أمّا المرحلة الثانية فقد بدأت في السبعينيَّات، واستمرت خلال الثمانينيَّات، وقامت على مبادئ نظرية التواصل؛ وكان سبب انتشار هذه النظرية هو الانتقادات التي تعرَّضت لها النظرية السلوكية، ذلك أنّ البرامج التي تقوم عليها النظرية السلوكية تعتمد على التكرار؛ وهي بذلك تفتقد عامل التواصل، حيث تقوم نظرية التواصل على استخدام الطالب للغة في أغراض واقعية؛ ويتم تقييم الطالب بناءً على إعطائه الإجابة، وليس من خلال الأخطاء التي يرتكبها [٢]. وقد تم تطوير العديد من البرامج التي تعتمد على هذه النظرية في التعليم، وهي تُعطي شيئا من التحكّم والحرية أثناء التعلّم.

ولمّا تعرّضت البرامج التي تقوم على نظرية التواصل للانتقاد بسبب عدم وجود نظام واضح وفاعل لاستخدام الحاسب الآلي في تطوير برامج تعليمية حديثة معتمدة يمكن أنْ تحل محلَّ البرامج التقليدية ظهرت برامج تقوم على عنصر التفاعل بين الطالب والمادة العلمية المبرمجة على الحاسوب، وبدأ ظهور تطبيقات ذكية وتفاعلية لتعليم اللغات توظف تقنيات المعالجة الآلية للغات لتوفير بيئة تعليمية تفاعلية تماثل المعلم في

الفصل الدراسيّ. وبعض هذه التطبيقات الشهيرة مثل المصحح الآليّ لأخطاء الإملاء والمصحح اللغويّ للأخطاء النحوية التي يستخدمها بالفعل كثير منَّا في برامج الكتابة مثل برنامج (MS Word) الشهير.

تُعتَبَر برمجيات تعليم اللغات عنصرًا مُهمًّا ومكملاً للمعلم في الفصل الدراسيّ. فالدارس يستطيع استخدام هذه البرمجيات للتدريب على مهارات اللغة المختلفة لعدد غير محدود من ساعات التدريب في أيّ وقت ومكان ملائم له [١١، ١٢]. وتتركز استخدامات تقنيات اللغة في التعليم في ثلاثة تطبيقات أساسية، هي: استخدام تقنيات معالجة اللُّغات الطبيعية (Natural Language Processing) في تعليم وتركيب المفردات والجمل الصحيحة لغويًّا، واستخدام تقنية التعرف الآليّ على الكلام (Automatic Speech Recognition) في تعليم مهارة النطق، واستخدام تقنية التعرف الآليّ على الكتابة (Automatic Handwriting Recognition) في تعليم مهارة الكتابة. وفيما يلي نوضح كيفية توظيف هذه التقنيات في مجال تعلم اللغة.

١- تقنيات معالجة اللغات الطبيعية (NLP)

تساعد تقنيات معالجة اللغات الطبيعية في مجال تعلم اللغة الطلاب على التعرف على التركيب الصحيح للمفردات اللغوية وقواعد بناء الجمل في تركيب سليم طبقًا لقواعد النحو والصرف.

ففي خطوات التعلم الأولية لمفردات اللغة العربية يمكن استخدام المدقق الآليّ لأخطاء الإملاء لاكتشاف أخطاء الإملاء في كتابة الطالب وتصحيحها، مثل:

مدرسه	والصَّواب: مدرسة
صوط	والصَّواب: صوت
أبراهيم	والصَّواب: إبراهيم
زبابة	والصَّواب: ذبابة

وهناك أساليب متعددة لبناء المدقق الإملائيّ، أشهرها هو استخدام قاموس مرجعيّ لمفردات اللغة مع بناء نموذج للأخطاء المحتملة.

ونظرًا لتميز اللغة العربية بخاصية الاشتقاق الصرفيّ الغزير، يوجد في هــذه اللغــة عدد ضخم من المفردات مما يزيد من صعــوبة استيعـاب الطالب لهذا الكــم مــن المفردات وخاصة في المراحل الأولية من تعلم اللغة. لذلك يمكن استخدام تقنيــة معالــج الصرف الآليّ لمساعدة الطالــب في التــعرف علــى قواعـــد الصرف للغــة العربيــة وتطبيقها على العديد من الأمثلة. ويمكن أيضًا استخدام معالج الصــرف الآلــيّ لتحليل المفردات التي يكتبها الطالب والتأكد من صحتها طبقًا لقواعد اللغة وتصحيح أخطائه وتدريبه على استخدام هذه القواعد الصرفية.

وعلى مستوى بناء الجمل يمكن استخدام المحلل النحويّ الآليّ لتحليــل البناء النحويّ للجمل التي ينشئها الطالب وتحكيم صحتها طبقًا لقواعد النحو في اللغة العربية وتصحيح الأخطاء أو اقتراح عدد من الحلول في حالة توفرها كما في هذه الأمثلة:

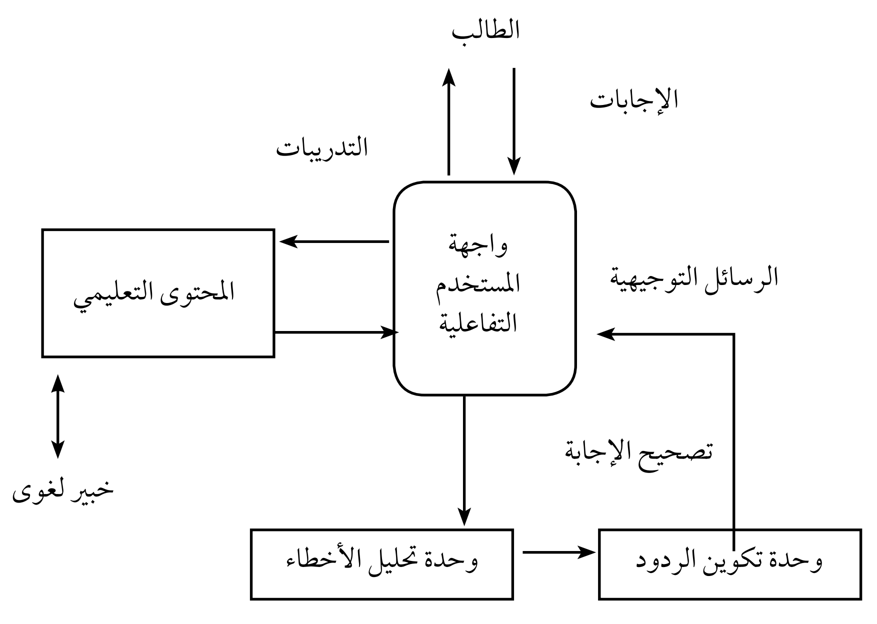
الحديقة جميلة	جملة اسمية من مبتدأ وخبر
الحديقة الجميلة	جملة غير تامة صفة وموصوف
ضرب محمد عليًّا	جملة فعلية من فعل وفاعل ومفعول به، ويلاحظ وجوب ظهور التنوين
الشجرة جميل	عدم تطابق النوع، والصَّواب: جميلة
الأولاد يلعب	عدم تطابق العدد، والصَّواب يلعبون
يلعب الأولاد	جملة صحيحة

تتميز اللغة العربية بخاصية المرونة النحوية بفعل عمليات التقديم والتأخير والحذف والإبدال والإضمار مما يمثل تحديًا لمتعلمي قواعد النحو في اللغة العربية. ويوفر المحلل النحويّ الآليّ أداة تعليمية فعالة للمساعدة في تعلم كيفية بناء الجمل صحيحة طبقًا لقواعد النحو.

كما تساعد أدوات الترجمة الآلية متعلمي اللغة الثانية كثيرًا، حيث توفر بعض الأدوات البسيطة - مثل ذاكرة الترجمة - وسيلة فعالة لمعرفة المفردات والنصوص اللغوية والمقابل لها باللغة الأخرى. ومما لا شك فيه أن توفر خدمة الترجمة الآلية المجانية على بعض مواقع شبكة الإنترنت قد ساهم بدرجة كبيرة في انتشار تعلم اللغات الأجنبية على الرغم من تدني مستوى الدقة الحالي لبرامج الترجمة الآلية، ولكنها تتطور

بسرعة كبيرة مع توفر النصوص المترجمة يدويًّا واستخدامها لتحسين نماذج الترجمة في هذه البرامج.

توجد طريقتان أساسيتان لاختبار المستخدم في أنظمة تعليم اللغات باستخدام الحاسب؛ الطريقة الأولى هي أسئلة الاختيار من متعدد، والطريقة الأخرى هي الأسئلة الإنشائية حيث يترك للمستخدم حرية استخدام مهاراته اللغوية في كتابة الإجابة المطلوبة كما في نظام (Arabic ICALL) الذي تم تطويره في كلية الحاسبات جامعة القاهرة، حيث يوجه النظام المستخدم للتعرف على أخطائه والتوظيف الخاطئ للوحدات اللغوية. ويساعد هذا النظام المستخدم على الاستفادة من أخطائه وتعلم كيفية إجراء التعديلات اللازمة لتصحيح أخطائه؛ وفي هذا النظام يتم استخدام محلل أخطاء يستعين بأدوات معالجة اللغة العربية، مثل محلل صرفيّ ومحلل نحويّ للتعرف على أخطاء المستخدم اللغوية وتتبع التصحيحات المطلوبة للأخطاء. ويتبع النظام المنهج التعليميّ للغة العربية الذي يتم تدريسه في المدارس المصرية في المرحلة الابتدائية. الشكل (٣-١) يوضح مكونات نظام (Arabic ICALL)، ويُمَثِّل واجهة المستخدم التفاعلية والمحتوى التعليميّ ومحلل الأخطاء اللغوية ووحدة تكوين ردود النظام.



الشَّكل ٣-١ : الشكل التوضيحي لمكونات نظام (Arabic ICALL) لتعليم قواعد اللغة العربية.

ويتبنى نظام (Arabic ICALL) منهجًا قائمًا على استخدام القواعد في التعرف على الأخطاء اللغوية؛ ويتميز هذا المنهج بقدرته على توفير تحليل تفصيلي لأخطاء الطالب اللغوية، حيث يستطيع تمثيل قواعد البناء اللغويّ طبقًا لنظم المعرفة اللغوية، وهذه القواعد تؤدي وظيفة مزدوجة لترميز البناء اللغويّ السليم، وأيضًا تكويد الأخطاء اللغوية المتوقعة على مستوى الصرف والنحو والدلالة، وتكون مصاحبة بالردود التعليمية المناسبة لكل خطأ. هذا المنهج لا يضمن تغطية كل الأخطاء المحتملة، ويمكن تلافي ذلك بالتركيز على نطاق تعليميّ محدد، ولكنه يتميز بعدم احتياجه إلى قواعد بيانات ضخمة لتدريبه مقارنة بالأنظمة التي تتبنى المنهج الإحصائيّ للتعرف على الأخطاء اللغوية.

٢- تعلم النطق باستخدام تقنية التعرف الآليّ على الكلام

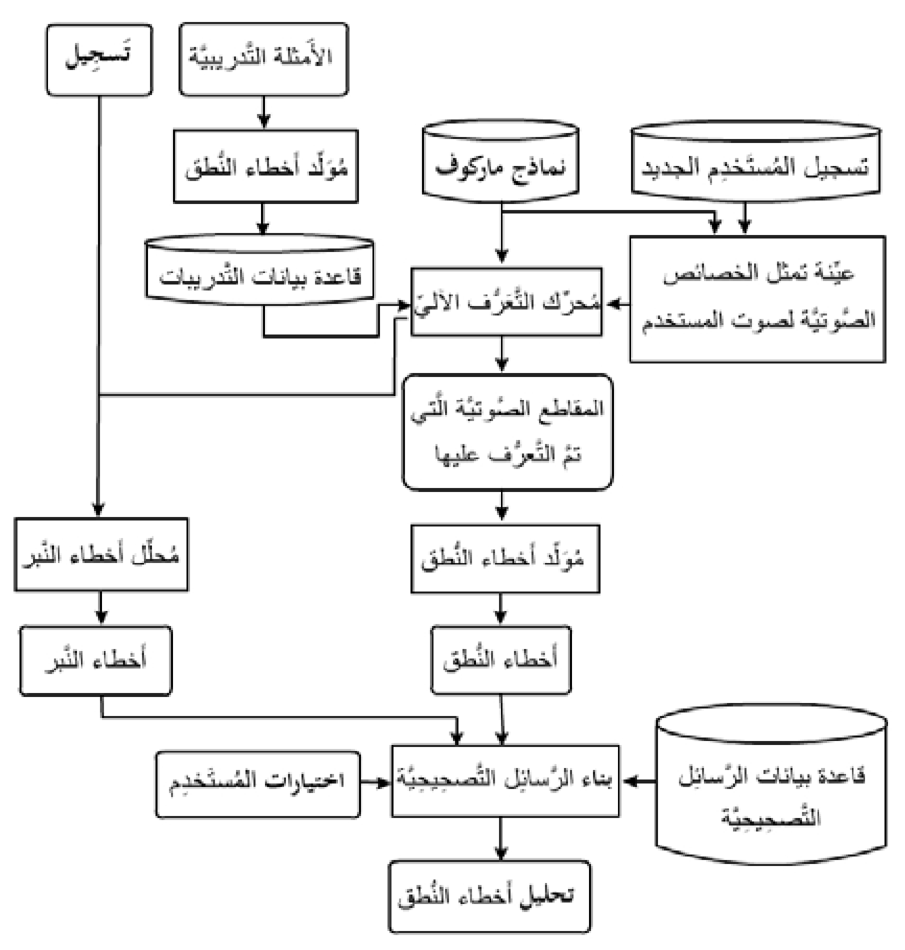
حازَ مجال تعليم النطق باستخدام الحاسب الآليّ على كثير من الاهتمام مؤخرًا. وقد تم إنجاز الكثير من الأبحاث في هذا المجال، لاسيما في أنظمة تعليم اللغات الأجنبية باستخدام الحاسب الآليّ.

يقوم الطالب -في هذه التدريبات- بالتدرُّب على قراءة حرف أو كلمة أو جملة معروضة أمامه على شاشة الجهاز. ويتم استخدام تقنية التعرف الآليّ على الكلام لتحكيم قراءة الطالب وتحديد ما إذا كانت قراءته صحيحة أم خاطئة.

وتُعَدُّ تقنية نماذج ماركوف المخفيَّة الأكثرَ استخدامًا في نُظُم تعليم القراءة. وتتكوَّنُ نماذج ماركوف المخفيَّة من عدد من النَّماذج؛ تُمِّثّلُ كلٌّ منها وحدةً من الوحدات الصوتيَّة. وتقومُ أنظمةُ تعليم القراءة بتقييم الجُملة الصَّوتيَّة الَّتي تمَّ تسجيلُها للمتعلِّم باستخدام هذه النَّماذج، حيث يتم تقطيع التسجيل الصوتيّ المدخَل من المستخدم إلى المقاطع الصوتية المختلفة طبقًا للكتابة الصوتية للمثال المستخدم في الاختبار. وتكون كلٌّ من هذه المقاطع الصوتية مصحوبةً بتقييم إحصائيّ يمثل درجة تماثل هذا المقطع الصوتيّ مع نماذج ماركوف الخاصة بالوحدة الصوتية لهذا المقطع. يتم تقييم قراءة المستخدم عن طريق حساب متوسط تقييمات المقاطع الصوتية المختلفة في العينة الصوتية تحت الاختبار؛ وإذا تخطت قيمة محددة (threshold) يتم اعتبار قراءة المستخدم قراءة

صحيحة. ويتمُّ اختيار القيمة المحددة (threshold) من التجارب المعيارية للنظام على عدد كبير من مستخدمي النظام.

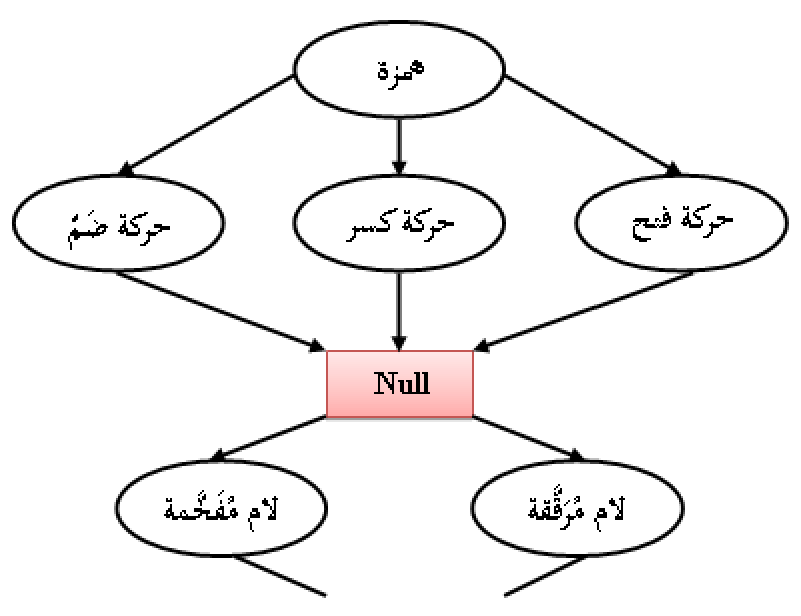
ويوضح الشكل (٣-٢) العناصر المكونة لنظام تعليم النطق.



الشَّكل ٣-٢: نظام تحليل أخطاء النطق.

يتم استخدام نماذج ماركوف في تطبيق تعليم القراءة والنطق بنفس طريقة توظيفها في تطبيق التعرف الآليّ على الكلام، مع اختلافٍ يتمثَّلُ في أنَّنا في هذا التطبيق نعُلَمِّ مسبقًا الجملة التدريبية التي ينطقها المستخدم مما يساعد على تحديد نطاق البحث لمحرك التعرف الآليّ؛ الأمر الَّذي يساعد على تحسين دقة النتائج. ويكون نطاق البحث فقط في

الصور المختلفة لأخطاء النطق في الجملة التدريبية المستخدمة، كما هو موضح بالشكل رقم (٣-٣).



الشَّكل ٣-٣: مثال لأخطاء النطق في جزء من كلمة.

يعتبر أسلوب تفاعل النظام مع المتعلم في التطبيقات التعليمية من أهم العناصر التي تؤثر في درجة استفادة المتعلم. لذلك يفضل - في تطبيقات تعليم القراءة - توفر درجة من الثقة في تقييم النظام لقراءة المستخدم لتقليل الحالات التي يصدر فيها تقييم خاطئ لقراءة الطالب؛ ويمكن في هذه التطبيقات حساب مُعامل الثقة؛ ويستخدم هذا المعامل في اختيار الرسالة المناسبة لتحكيم قراءة مستخدم النظام. هناك طرق متعددة تم اقتراحها لحساب هذا المعامل، من أفضلها معامل نسبة الأرجحيَّة (Likelihood Ratio) الذي يتم حسابه عن طريق ترتيب المقاطع الصوتية في قراءة المستخدم بناء على التقييم الإحصائيّ المصاحب لها؛ ثم يتم حساب معامل الثقة بالمعادلة التالية:

$$\text{معامل الثقة} = \frac{\text{متوسط تقييم ارجح مسار في نماذج ماركوف}}{\text{متوسط تقييم ثاني أرجح مسار في نماذج ماركوف}}$$

ويعتمد هذا المعامل على قيمة الاختلاف بين أرجح تقييم من البدائل المتاحة لنطق الكلمة موضع الاختبار والتقييم التالي في الترتيب؛ فوجود اختلاف كبير يدل على وجود درجة عالية من الثقة في تحكيم النظام لقراءة المستخدم.

وتعتمد الرسائل التصحيحية في هذا النظام على مدى ثقة النظام من النتائج (معامل الثقة)؛ وإذا كان معامل الثقة صغيرًا يسير النظام على إحدى هذه الطرق:

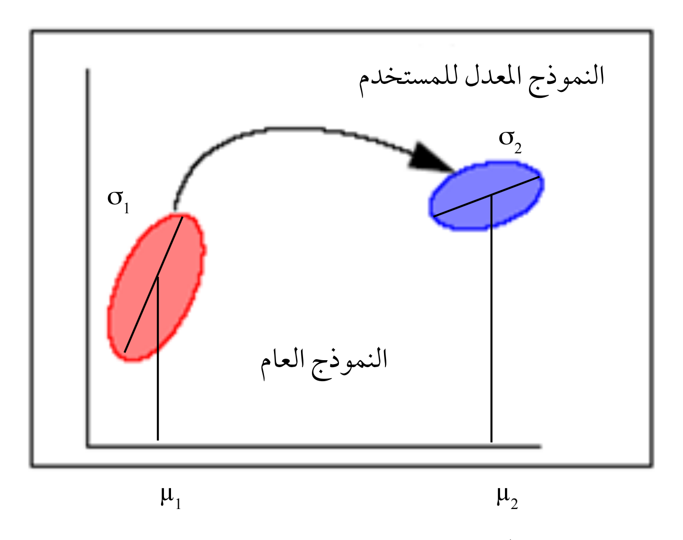
١- إهمال الخطأ تمامًا وعدم ظهور رسالة له (وهذا جيد للمبتدئين، لأن إظهار إنذار خاطئ يثبط من عزم المتعلم).

٢- طلب إعادة الجملة من المستخدم لأنها غير واضحة.

٣- إظهار رسالة توضح وجود خطأ غير محدد، وطلب إعادة الجملة.

٤- إظهار رسالة بأكثر الأخطاء شيوعًا في هذا الموضع.

ويمكن تحسين أداء أنظمة تعليم القراءة بدرجة كبيرة عن طريق تكييف وتحويل النماذج الصوتية (Models Adaptation) لتشابه خصائص صوت المستخدم. ومن أشهر طرق تحويل نماذج ماركوف طرق التحويل الخطيّ، نظرًا لسهولة تنفيذها وفاعلية نتائجها. وفي هذه العملية يتم تجميع عدد قليل من الجمل من المستخدم لاختيار أقرب حزمة صوتية لصوت المستخدم؛ وهذه الحزمة تستخدم كنموذج مرجعي لصوت المستخدم. ثم يُطلب من المستخدم إدخال عدد من الجمل يتم تحكيمها باستخدام هذا النموذج المرجعي. وإذا رأي النظام أن الجملة خالية من أخطاء القراءة يقوم بوضعها في المجموعة التي سيتم استخدامها في عملية تحويل النموذج الصوتي. تستمر هذه العملية حتى يتم تجميع كمية كافية من الجمل لبدء عملية تحويل النماذج الصوتية باستخدام مصفوفات التحويل الصوتيّ الخطيّ إلى شكل يحاكي صوت المستخدم. ويوضح الشكل رقم (٣-٤) رسمًا توضيحيًّا لعملية التحويل الخطيّ في نماذج ماركوف.



الشَّكل ٣-٤: عملية تحويل النماذج الصوتية.

ومن تطبيقات تعليم النطق التي تمثل درجة عالية من التحدي: استخدام الحاسب الآليّ في تعليم قراءة وتجويد القرآن الكريم. فبالمقارنة مع تطبيق تعليم اللغات - حيث توجد فروق واضحة بين النطق الصحيح والأخطاء - تحتاج بعض قواعد النطق في علم التجويد إلى درجة عالية من الحساسية من أجل الحكم بصحة تطبيق القاعدة التجويدية؛ مما يتطلب استخدام معالج آليّ للتعرف على الكلام ذي درجة عالية جدًّا من الدقة. والتحدي الأكبر هو معالجة هذه التقنيات للأحكام التجويدية والفونيمات فوق المقطعية التي ينبغي على دارس القرآن تعلمها.

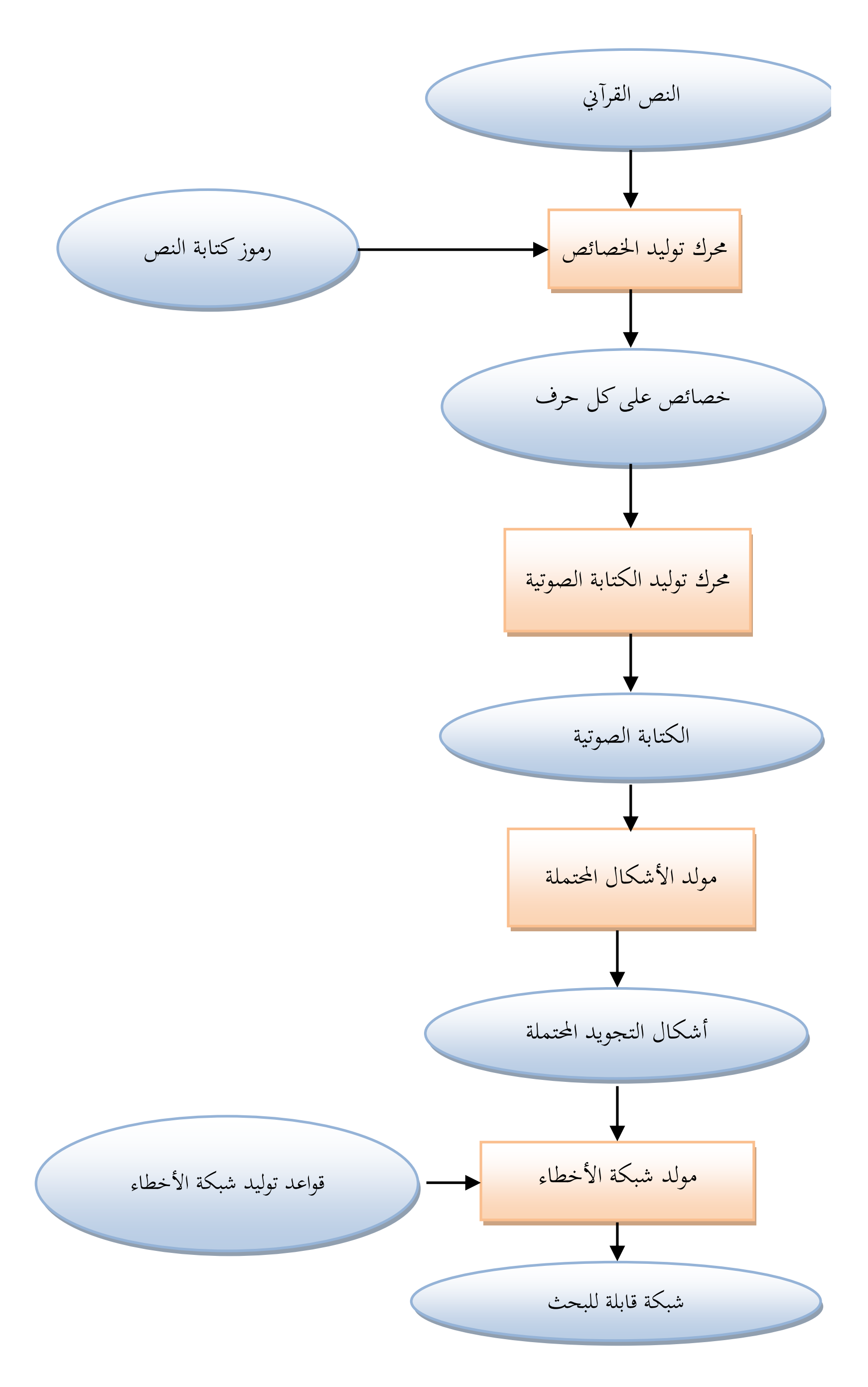
ويوضح الشكل (٣-٥) مثالاً لشاشة تطبيق تعليم التجويد حيث يستطيع المستخدم بدء التدرُّب باختيار الدرس الذي يريد التدرب عليه، فيقوم البرنامج بتوليد تدريبات خاصة لهذا الدرس، فيستمع المستخدم لتلاوة هذه التدريبات بصوت قارئ قياسيّ ثم عليه أن يكررها، فيقوم البرنامج بتوليد تقرير مكتوب ومنطوق عن قراءة المستخدم يشرح له فيها أخطاءه التجويدية واللغوية كما هو موضح بالشكل. ويظهر في التقرير المكتوب المقطع الذي قرأه المستخدم مع تلوين الحرف أو الحروف التي فيها الخطأ باللون الأحمر مع شرح للخطأ المرتَكَب؛ ويتم توليد التقرير المنطوق بحيث يظن المستخدم أن الحاسوب يفهمه ويتفاعل معه، لأنه يخبره بالخطأ صوتيًّا.

ويتضمن التقرير الصوتيّ مثالاً منطوقًا عن الحكـــم الذي كـــان فيه الخطأ من قبل المستخدم.



الشَّكل ٣-٥: شاشة التدريبات في برنامج تعلم التجويد.

ويوضح الشكل (٣-٦) نظامَ بناء الأخطاء التجويدية المستخدمة في هذا التطبيق؛ وهو يستخدم مجموعة من القواعد تم تصميمها بصورة مرنة تُسَهِّل التَّعديلَ بالإضافة أو الحذف للأخطاء التي يُحكِّمها النظام [٤].



الشَّكل ٣-٦: محرك توليد شبكة أخطاء النطق طبقًا لأحكام التجويد.

يعتمد توليد شبكة الأخطاء على محرك توليد الكتابة الصوتية للقرآن الكريم الذي يتكون من عدة طبقات؛ وهى محرك الأحداث الذي يبحث خلال الرسم العثماني للقرآن الكريم عن خصائص كل حرف ويولد له كودًا خاصًّا ويُبَيِّنُ حالة نطقه (منطوق أو غير منطوق، مشدد أو غير مشدد، التشكيل الخاص بالحرف ...) وخصائصه الصوتية (مجهور، مهموس، مخرج الحرف، ...). ويقوم محرك توليد الكتابة الصوتية بتحليل هذه الأكواد/ الشَّفرات وخصائص توليد الكتابة الصوتية الصَّحيحة طبقًا لأحكام تلاوة القرآن الكريم؛ ثم يقوم محرك مقارنة النماذج (القواعد) بتجميع كل هذه المعلومات وتوليد أخطاء التلاوة المتوقعة.

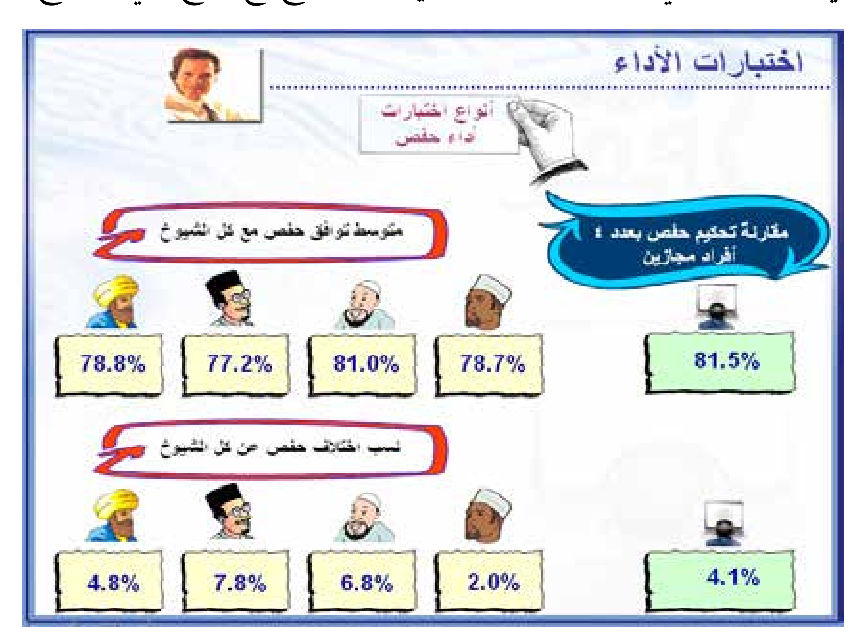
والقاعدة الآتية توضح طريقة توليد الأخطاء الافتراضية:

> بافتراض أنَّ الوحدة الحالية:
> فونيم= Ph_c، نوع الحركة= V_c، طول الحركة= Lc،
> مشدد= S_c، منطوق= P_c، إدغام= E_c
> والوحدة السابقة:
> فونيم= Ph_p، نوع الحركة= V_p، طول الحركة= L_p،
> مشدد= S_p، منطوق= $\pm P$
> والوحدة التالية:
> فونيم= Ph_n، نوع الحركة= V_n، طول الحركة= L_n،
> مشدد= S_n، منطوق= P_n
> إذًا أضف مسارًا لخطأ التلاوة بالخصائص التالية:
> رقم الخطأ=C، نوع الخطأ=T، كلمة الخطأ=W، مقدار شيوع الخطأ=F، فونيم=Ph، نوع الحركة=V، طول الحركة=L، مشدد=S، منطوق=P، مدغم=E

وتستخدم هذه النماذج (القواعد) بعد توليدها للمقارنة مع الكتابة الصوتية للقرآن الكريم لتوليد شبكة مسارات الأخطاء المتوقعة. ويقوم مولد شبكة مسارات الأخطاء بترتيب النماذج (القواعد) التي تطابقت بشكل تنازلي حسب مدى تطابقها مع الحالة الحالية ثم يهمل كل النماذج (القواعد) التي تولد نفس الخطأ ما عدا أولها، ثم في النهاية تولد الشبكة بشكل يناسب نظام التعرف على الصوت. والوحدة المستخدمة في بناء شبكة الأخطاء في هذا الاختراع شبيهة بتلك التي تستعمل في طرق تعليم تلاوة

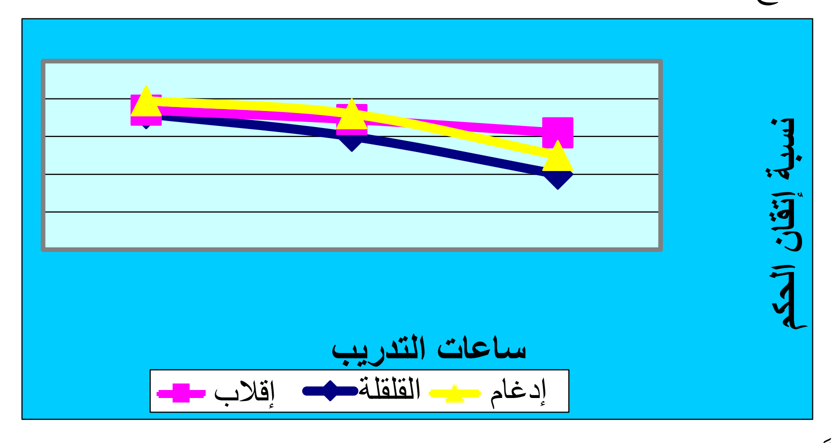
القرآن الكريم، حيث إن مفهوم الفونيم غير ملائم للمستخدم. وتتكون الوحدات في هذا النظام من حرف+حركة قصيرة، حرف+حركة طويلة، حرف ساكن، حرف مشدد+حركة قصيرة، حرف مشدد+حركة طويلة، حرف مشدد ساكن.

تعد اختبارات الأداء العملية من أهم وسائل تقييم أنظمة تعلم القراءة لبيان دقتها وأثرها على منحنى التعلم لدى المستخدم المبتدئ. وأحد هذه الاختبارات هو تقييم أداء النظام بالمقارنة مع أداء معلم بشريّ باستخدام نفس المعطيات. ففي تجربة لقياس الاختلاف البشري المسموح به في تعليم التجويد عن طريق التحقق من نسبة الاتفاق والاختلاف في تحكيم الشيوخ المجازين لقراءة المتعلم العادي، تم تسجيل ثلاثمئة محاولة لنطق كلمات أو جمل تحتوي على أحكام تجويد محددة من شخص عادي، ثم طلب من أربعة شيوخ مجازين تحكيم هذه المحاولات الثلاثمئة، كلٍّ على حدة، ثم تم إدخال نفس المحاولات على برنامج تعليم التجويد لتحكيمها. أظهرت النتائج أن: متوسط اتفاق تحكيم أي محكم مجاز مع أي محكم مجاز آخر في حدود ٨٠٪، وهي نفس نسبة اتفاق البرنامج مع أيٍّ من الشيوخ المجازين. والنسبة التي خالف فيها البرنامج إجماعهم حوالي ٤٪ فقط، وهي نفس نسبة اختلاف أي من الشيوخ مع إجماع باقي الشيوخ.



الشَّكل ٣-٧: توضيح لمدى اتفاق واختلاف المشايخ مع بعضهم البعض ومع برنامج تعليم التجويد.

وفي اختبار آخر لقياس أثر برنامج تعليم التجويد على تعلم الفرد المبتدئ لحكم أو أكثر من أحكام التجويد، تمت الاستعانة بمجموعة من الأفراد العاديين الذين لا يجيدون أحكام قراءة القرآن الكريم، ثم إجراء اختبار قبلي - قبل استخدام التقنية - لتحديد مستوى المستخدم، ثم إتاحة الفرصة له للتعلم على الحكم أو القاعدة المطلوبة مع سماع أمثلة ونماذج صوتية للحكم، وتم التركيز على قواعد ثلاث، هي: القلقلة، وحكم إدغام النون الساكنة والتنوين بغنة، وحكم إقلاب النون الساكنة والتنوين. وتم اختبار المستخدم مرة أخرى بعد فترة من التدريب - ساعة تقريبًا - باستخدام البرنامج، ثم مرة ثالثة بعد ساعة تدريب أخرى. وأظهرت النتائج ارتفاع منحنى التعلم لمجموعة المستخدمين بنسب كبيرة حيث بدأ بنسب تقترب من ٤٠٪ وارتفع إلى ما يزيد على ٧٠٪ خلال بضع ساعات من التدريب باستخدام البرنامج . وفي هذا برهان على فعالية التقنية وكفاءتها. ويوضح الشكل (٣-٨) متوسط نتائج المشاهدات التي أثمرتها التجربة مع عشرة مستخدمين.



الشَّكل ٣-٨: شكل بياني لقياس أثر تقنية تعليم التجويد في تطوير منحنى التعلم لدى المبتدئين.

٣- تعلم الكتابة باستخدام تقنية التعرف الآليّ على الكتابة

مع توفر أجهزة الكمبيوتر اللوحية، تم مؤخَّرًا تطوير تطبيقات لاستخدامها في تعليم مهارة الكتابة باستخدام تقنيات التعرف الآليّ على الكتابة، حيث توفر هذه الأجهزة شاشات تفاعلية يمكن الكتابة على سطحها. ويقوم التطبيق بتحكيم درجة جودة كتابة الطالب طبقًا لقواعد كتابة الخط العربيّ.

ولقد قام فريق بحثيٌّ في كلية الحاسبات بجامعة القاهرة بتطوير تطبيق لتعليم كتابة الخط العربيّ للأطفال في المراحل التعليميَّة الأوَّليَّة. ويساعدُ هذا التطبيقُ الطِّفلَ على تحسين جودة خطه باستخدام مستويين من التَّدريبات بصورةٍ مشابهة لتدرُّج تعليم الكتابة في دُرُوس الخطّ العربيّ في مناهج التعليم للمرحلة الابتدائية؛ المستوى الأول هو الكتابة الموجهة للدروس الأولية لتعليم الكتابة باستخدام صور منقطة لأشكال الحروف المختلفة؛ أمَّا المستوى الثاني فهو الكتابة الحرة، حيث يمارس الطالب التدريب على الكتابة بصورة غير مقيدة.

في تدريبات المستوى الأول، يقوم الطالب بالتدرُّب على كتابة حرف أو كلمة أو جملة معروضة أمامه على شاشة الجهاز، حيث يتم عرض صورة متحركة توضح طريقة الكتابة المثالية للنموذج الموضح على الشاشة للمستخدم، ثم يتم عرض صورة لهذا النموذج على الشاشة بلون شفاف. ويُطلَبُ من المستخدم الكتابة على هذا النموذج عن طريق المرور فوقه بالقلم بنفس طريقة الكتابة المثالية السابق عرضها له؛ ويحتوي هذا النموذج الشفاف على عدد من نقاط التحكيم غير المرئية للمستخدم كما هو موضح بالشكل (٣-٩). ثُمَّ تُستَخدمُ هذه النِّقاط لتحكيم العناصر التالية في كتابة المستخدم:

١- درجة قرب كتابة المستخدم من موضع نقاط التحكيم.

٢- ترتيب مرور كتابة المستخدم على نقاط التحكيم.

٣- عدد الوقفات في كتابة المستخدم.

يتم تجميع هذه التقييمات في تقييم إجماليّ يستخدم لإنتاج رسائل تصحيحية مناسبة للأخطاء المحددة، حيث تساعد هذه الرسائل في إرشاد المستخدم لطريقة الكتابة السليمة وتوجيه حركة اليد في الاتجاه الصحيح والتحكم في القلم المستخدم، كما تساعده على الكتابة بطريقة سهلة التعلم وواضحة.



الشَّكل ٣-٩: نموذج تعلم الكتابة الموَجَّه.

بعد أن يجتاز الطالب المستوى الأول من التدريبات يمكن أن ينتقل إلى المستوى الثَّاني في الكتابة الحرة، حيث يكون قد تمكن من إتقان المهارات الأساسية لكتابة الحروف؛ لذلك يتم تدريب الطالب – في المستوى الثاني – على الكتابة بدرجة أكبر من الحرية بدون توجيهٍ لاختبار مدى استيعاب الجهاز الحركي عند الطالب لأشكال كتابة الحروف وتخزينها في الذاكرة الباطنية.

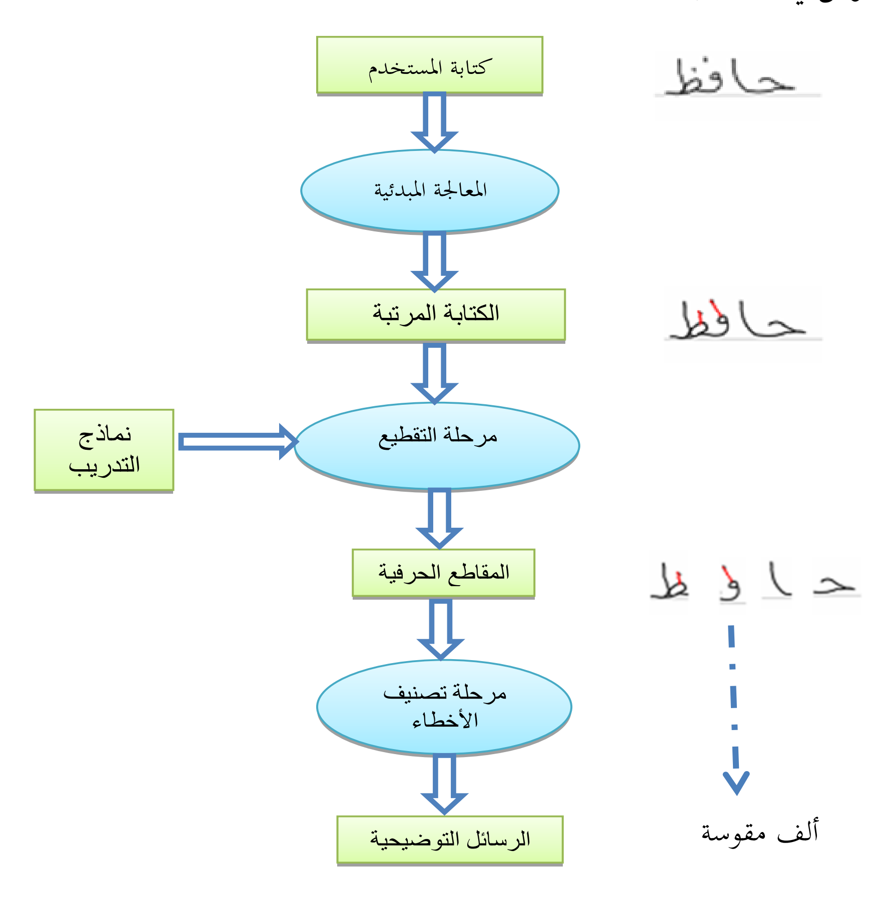
وفي تدريبات هذا المستوى يتم عرض صورة متحركة توضح طريقة الكتابة المثالية للنموذج الموضح على الشاشة للمستخدم، ثم يُطلَبُ من المستخدم كتابة النموذج بنفسه على مساحة بيضاء. ويتم تحكيم جودة كتابة المستخدم، ثُمَّ تُعرَضُ نتيجة تقييم هذه الكتابة مع توضيح نوعية الخطأ وموضعه في أي حرف من حروف النموذج المستخدم. ولعمل هذا التحكيم يتم تنفيذ مرحلتين من المعالجة لكتابة المستخدم؛ حيثُ يتم في المرحلة الأولى تحديد المقاطع الحرفية في كتابة المستخدم وحدود البداية والنهاية لكل حرف في عينة الكتابة، وذلك باستخدام نماذج ماركوف المخفية الَّتي سبق شرحها في تطبيق تعليم القراءة.

وتُستَخدَمُ هذه النماذج نظرًا للتطابق الكبير بين عمليتي النطق والكتابة؛ فالأولى هي تتابع من الأصوات المنطوقة، والثانية هي تتابع من الحروف المكتوبة، مع اختلاف بسيط في الثانية، إذ عادة تُضاف النقط والعلامات التشكيلية بعد إتمام كتابة الكلمات. وهذا التحرك الخلفي في اتجاه الكتابة يُسبب نوعًا من التعارض مع الفرض الأساسي لنماذج ماركوف المخفية من نوع (Ergodic HMM)؛ وهو أنها تستطيع عمل نمذجة للبيانات المتتابعة في اتجاه أمامي فقط.

وللتغلب على هذا التعارض، أضِيفَت مرحلة قبل المعالجة لإعادة ترتيب تتابع الوحدات المكتوبة في عينة الكتابة لتكون في تسلسل أمامي، مما يسهل استخدام نماذج ماركوف المخفية من نوع (Ergodic HMM) لعمل نمذجة لهذه البيانات.

وبعد تحديد المقاطع الحرفية في كتابة المستخدم، يتم تحليل جودة كتابة كل مقطع وتحديد نوع الخطأ في شكل كتابة الحرف إن وجد، ويتم ذلك باستخدام عدد من المصنفات الثنائية، يتخصص كل مصنف منها في إعطاء قرار ثنائي عن وجود خطأ محدَّد في كتابة الحرف أو عدم وجوده.

في الشكل (٣-١٠) يتم عرض مُخَطَّط لمراحل تحكيم الكتابة الحرة للمستخدم، وكما نرى في مراحل تحكيم كلمة «حافظ» بعد المعالجة المبدئية وإعادة ترتيب اتجاه الكتابة في الكلمة تم تحديد المقاطع الحرفية في الكلمة، وهي هنا أربعة مقاطع؛ وتم تحكيم جودة كتابة كل حرف والأخطاء الملحوظة فيه؛ مثل حرف الألف، وهي مكتوبة بشكل مقوس في هذه العينة.



الشَّكل ٣-١٠: مراحل تحكيم الكتابة الحرة.

ولتصنيف أخطاء الكتابة هناك العديد من المصنفات الآلية التي يمكن استخدامها. وتُعَدُّ آليَّات المتَّجهات الدَّاعمة (support vector machine) من أفضل الوسائل في تطبيقات التصنيف الثنائية، مثل تطبيق تعليم الكتابة، حيث يتم تصنيف كل حرف مكتوب باحتوائه على خطأ محدد أم لا. والجدول (٣-١) يوضح أمثلة لعدد من أخطاء

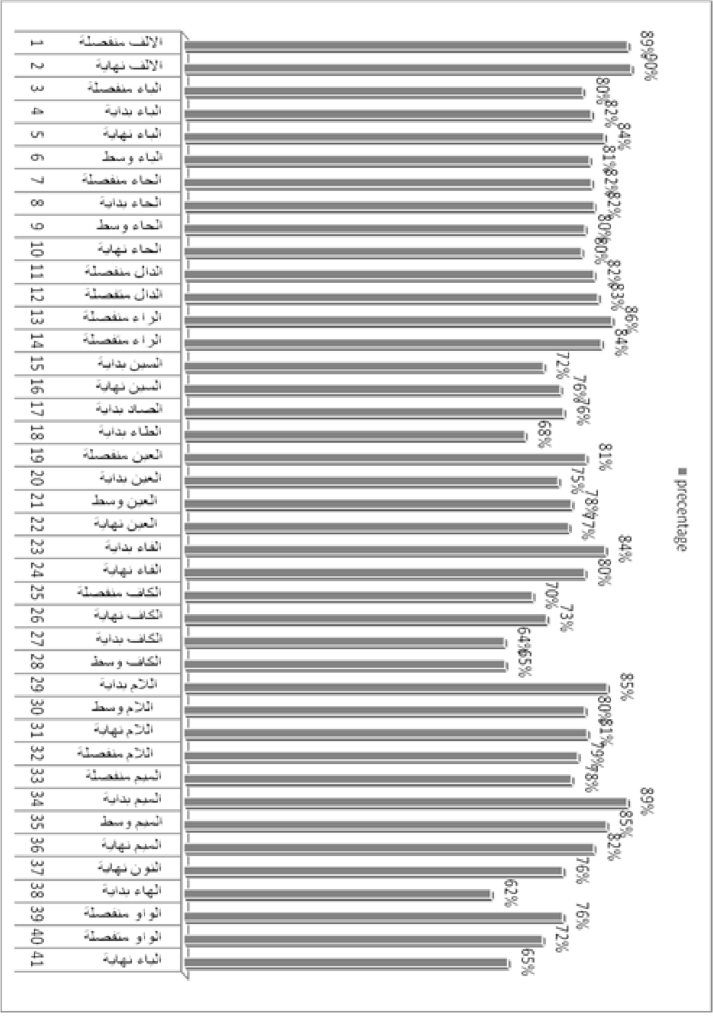
الكتابة في عدد من الحروف، حيث يتم بناء عدد من المصنفات الثنائية مساو لعدد الأخطاء التي تم حصرها لكل حرف. ويتم تدريب هذه المصنفات باستخدام عينات من الكتابة المثالية وعينات من الكتابة بالخطأ المحدد. ويتم استخدام جزء من عينات الكتابة لتوليف أفضل اختيار لبارامترات المصنف.

عينة خطأ	نوع الخطأ	عينة جيدة
	تدويرة حرف الواو صغيرة	
	أربع نبرات في حرف السين	
	تدويرة حرف السين غير مكتملة	
	حرف الراء شبه الدال	
	لا يوجد انحناء في رسم الراء	
	لا يوجد انحناء في رسم الحاء	
	حرف الصاد بدون نبرة	
	حرف العين يُشبه رقم أربعة	
	لا يوجد عنق لحرف الفاء	
	تدويرة حرف الهاء مفتوحة	

الجدول ٣-١ : نماذج لأخطاء كتابة بعض الحروف.

تم اختبار هذا التطبيق على طُلاَّب عدد من مدارس المرحلة الابتدائية في المرحلة العمرية ٦-١١ سنة. تم تدريب عدد ٥٠ طالب على استخدام التطبيق، ثم استخدم الطلاب التطبيق للتدرب على كتابة الأمثلة المتاحة. بعد عدد قليل من الحصص التعليمية تمكنت نسبة كبيرة من الطلاب من تحسين كتابتهم؛ وقد تمكن بعضهم من تقليد طريقة الكتابة المثالية للأمثلة. وأشارَ تحليل نتائج التجربة إلى وجود تفاوت في قدرة التطبيق

على تتبع أخطاء الكتابة المختلفة كما هو موضح بالشكل (٣-١١) حيث تراوحت الدقة للحروف المختلفة ٦٠٪ - ٩٠٪ طبقًا لدرجة الصعوبة في شكل الحرف وتعرُّف التطبيق على نوع الخطأ في كتابته. وبصورة ما وُجِدَ أن درجة استفادة الطلاب الأصغر سنًّا من النظام أكثر من قُرنائِهم الأكبر سنًّا؛ وهذه نتيجة متوقعة إلى حد كبير بسبب مرونة جهاز الكتابة عند الصغار، مما يسهل تعلم التغيير في طريقة كتابة الحروف وتعديل هذه الأشكال في العقل الباطن للطفل، مما يسهل عليه استخدامها بصورة آلية في المستقبل.



الشَّكل ٣-١١: نسبة الدقة في التعرف على أخطاء كتابة الحروف المختلفة.

٤- مقترحات بَحثِيَّة

تُعَدُّ مشاركة الباحثين اللُّغويين في تطوير استخدام تقنيات معالجة اللغة العربية في تطبيقات التعليم من الأمور المهمَّة. فعلى سبيل المثال، لتصميم برامج فعالة لتعليم النطق، ينبغي توفُّر تحليل لغوي دقيق لأنواع أخطاء النطق وأسبابها وكيفية معالجتها. ويُعَدُّ الباحثون اللُّغويُّون أفضلَ مَن يؤدِّي هذه المهمة. فجدير بالذكر أن علم الصوتيات في اللغة العربية يُعَدُّ من أقدم العلوم مقارنة باللغات الأخرى، حيث تم وضع أسسه على يد العالم الجليل الخليل بن أحمد الفراهِيدِيّ في القرن الثاني الهجريّ. وحتى الآن تم إنتاج تراث ضخم من الأبحاث في الخصائص الصوتية للغة العربية والدراسات المقارنة مع اللغات الأخرى. وبدون شك فإن توظيف هذه الأبحاث في تطوير تقنيات تعلم النطق باستخدام الحاسب سوف يُسهم في تطوُّر هذه التقنيات والحصول على درجة عالية من الدقة تقارب المعلم الحقيقي.

وتستطيع الأبحاث اللغوية أيضًا توفير التحليل اللغويّ لأخطاء الكتابة وتصنيفها وتعيين معدل تكرارها وتحديد مرجعيتها طبقًا لقواعد البناء اللغويّ في اللغة العربية. فكل هذه الخصائص تساعد على تصميم برامج تعليمية للغة العربية بصورة ميسرة ومفيدة للدارسين حيث يتم التركيز فيها على العناصر الفعالة التي تقود المتعلم لفهم واستيعاب قواعد اللغة من ناحية، وتوظيفها في كتابة النصوص بلغة عربية سليمة من ناحيةٍ أخرى.

ومن ناحيةٍ أخرى، يستطيع الباحثون المعنيُّون بالخط العربيّ توفير قواعد تعليمية لتحسين طرق الكتابة وإظهار جماليات الخط العربي. ويمكن استنتاج هذه القواعد عن طريق تحليل مُشكلات الكتابة في عينات من الخط تمثل نماذج الكتابة بدرجات إتقان متفاوتة.

ببليوجرافيا مرجعيَّة

١. بشر (كمال محمد): علم اللغة العام (الأصوات)، دار غريب، ط٢، ١٩٧١م.

٢. الحصري (محمود خليل): أحكام قراءة القرآن الكريم، مكتبة السنة، ط١، ٢٠٠٢م.

٣. الصفاقسي (أبو الحسن علي بن محمد النوري): تنبيه الغافلين وإرشاد الجاهلين عما يقع لهم من الخطأ حال تلاوتهم لكتاب الله المبين، المطبعة الرسمية للجمهورية التونسية، ١٩٧٤م.

4. Abdallah, M.; Al-Marri, M.; Abdou , S.; Rashwan, M.; El-Gamal, M. (2015). "Improving Holy Qur'an recitation system using Hybrid Deep Neural Network-Hidden Markov Model approach", Third International Conference on Islamic Applications in Computer Science And Technology, 1-3, Turkey.

5. Abdou, S.; Elgammal, A.; Fahmy, A. (2010). "A Tool for Arabic Handwriting Training". The Fifth Conference of Learning International Networks Consortium (LINC), 23-26 May 2010, MIT, Cambridge, Massachusetts, USA.

6. Abdou, S.; Hamid, S.; Rashwan, M.; Samir, A.; Abdel-Hamid, O.; Shahin, M.; Nazih, W. (2006). "Computer Aided Pronunciation Learning System Using Speech Recognition Techniques", INTERSPEECH 2006 - ICSLP, Pittsburgh, PA, USA.

7. Abdou, S.; Rashwan, M.; Al-Barhamtoshy, H.; Jambi, K. and Al-Jedaibi, W. (2014). "Speak Correct: A Computer Aided Pronunci-ation Training System for Native Arabic Learners of English" Life Science Journal 2014; 11(10) pp. 370-380.

8. Al-Barhamtoshy, H.; Abdou, S.; Rashwan, M. (2014). "Mobile Technology for Illiterate Education". Life Science Journal 2014; 11(9) pp. 242-248.

9. Al-Barhamtoshy, H.; Alwajih, F.; Abdou, S. (2012). "A Toolkit For Teaching Arabic Handwriting", International Journal of Computer Applications, Vol. 49, No. 23.

10. Al-Marri, M., Raafat, H., Abdallah, M., Abdou, S., & Rashwan, M. (2018). Computer Aided Qur'an Pronunciation using DNN. Journal of Intelligent & Fuzzy Systems, vol. 34, no. 5, pp. 3257-3271, 2018
11. Bacha, K.; Jemni, M.; Zrigui, M. (2016). "Towards a Learning System Based on Arabic NLP Tools". International Journal of Information Retrieval Research (IJIRR) 6.4.
12. Bax, S. (2003). "CALL - past, present and future", System 31, 1: 13-28.
13. Bax, S.; Chambers, A. (2006). "Making CALL work: towards nor-malisation", System 34, 4: 465-479.
14. Cushion, S.; Hémard, D. (2003). Designing a CALL Package for Arabic While Learning the Language Ab Initio. Computer Assisted Language Learning (CALL): An International Journal, Vol. 16, No. 2-3, 259-266, Belgium: SWETS & ZEITLINGER publisher.
15. Elaraby, Mohamed S., et al. "A Deep Neural Networks (DNN) Based Models for a Computer Aided Pronunciation Learning System." International Conference on Speech and Computer. Springer International Publishing, 2016
16. El-Kasasy, M. S. (1992). "An Automatic Speech Verification System", Ph.D. Thesis, Cairo University, Faculty of Engineering, Department of Electronics and Communications, Egypt.
17. Franco, H. et al. (2000). "The SRI EduSpeak system: Recognition and pronunciation scoring for language learning", Proc. of InSTIL, Scotland, 123-128.
18. Fuerstenberg, G. (1993). A la rencontre de Philippe: Videodisc, Software, Teacher's Manual and Student Activities Workbook: Yale University Press [Online]: http://web.mit.edu/fll/www/projects/Philippe.html
19. Hagen A.; Pellom, B. (2005). "Data driven sub-word unit modeling for speech recognition and its application to interactive reading tutors," in Proc. European Conference on Speech Communication and Technology, Lisbon, Portugal, Sept. 2005, pp. 2757–2760.

20. Hager Morsy, M. S., Aljohani, N., Shoman, M., & Abdou, S. (2018). Automatic Speech Attribute Detection of Arabic Language. International Journal of Applied Engineering Research, 13(8), 5633-5639.

21. Hamid S.; Rashwan, M. (2004). "Automatic Generation of Hypotheses for Automatic Diagnosis of Pronunciation Errors" Proceedings of NEMLAR International Conference on Arabic Language Resources and Tools, pp. 135-139, Cairo, Egypt.

22. Hamid, S. (2005). "Computer Aided Pronunciation Learning System using Statistical Based Automatic Speech Recognition". PhD thesis, Cairo University, Cairo, Egypt.

23. Hammadi, M.; Bezine, H.; Njah, S.; Alimi, A. M. (2012): "Towards educational tool for Arabic Handwriting Learning", International Conference on Education and e-Learning Innovations.

24. Hosseny, I.; Abdou, S.; Fahmy, A. (2011). Using Advanced Hidden Markov Models for Online Arabic handwriting recognition", ACPR2011, Bekeen, China, 29-30.

25. Hubbard, P.; Levy, M. (2016). "Theory in computer-assisted language learning research and practice". The Routledge Handbook of Language Learning and Technology (2016): 24.

26. Kosaka, T.; Sagayama, S. (1994). "Tree–structured speaker clustering for fast speaker adaptation", proceedings of the 1994 IEEE International Conference on Acoustics, Speech and Signal Processing 1, 245–248. IEEE, New York.

27. Magdy, M. Shalaan, K.; Fahmy, A. (2010). Morphological Analysis of Ill-formed Arabic Verbs in Intelligent Language Tutoring Framework. In FLAIRS-23, The 23rd International FLAIRS Conference, Florida, USA

28. Mars, A.; Antoniadis, G. (2015). "Handwriting recognition system for Arabic language learning". Information Technology and Computer Applications Congress (WCITCA), 2015 World Congress on. IEEE.

29. Marse, J. F. et al, (1991). Handwriting Training: Computer-Aided Tools for Premedial Teaching, Development of Graphic Skills. J. Wann, A.M. Wing, N. Suvik, eds., pp. 249-258, Academic Press.
30. Mote, N.; Johnson, L.; Sethy, A.; Silva, J.; Narayanan, S. (2004). Tactical Language Detection and Modeling of Learner Speech Errors: The case of Arabic tactical language training for American English speakers. In the Proceedings of InSTIL/ICALL2004 – NLP and Speech Technologies in Advanced Language Learning Systems, Venice. at http://sisley.cgm.unive.it/ICALL2004/link13.htm
31. Samir, A.; Abdou, S. M.; Khalil, A. H.; Rashwan, M. (2007). "En-hancing usability of CAPL system for Qur'an recitation learning", INTERSPEECH 2007 - ICSLP, Antwerp, Belgium.
32. Schneider, E.W.; Bennion, J.L. (1984). "Veni, vidi, vici, via videodisc: a simulator for instructional courseware". In Wyatt D.H. (ed.) Computer-assisted language instruction, Oxford: Pergamon.
33. Underwood, J. (1984). Linguistics, computers and the language teacher: a communicative approach, Rowley, Massachusetts: Newbury House.
34. Warschauer, M. (1996). "Computer-assisted language learning: an in-troduction". In Fotos S. (ed.) Multimedia language teaching, Tokyo: Logos International [Online]: http://www.ict4lt.org/en/warschauer.htm.
35. Warschauer, M. (2000). "CALL for the 21st Century", IATEFL and ESADE Conference, 2 July 2000, Barcelona, Spain.
36. Warschauer, M.; Healey, D. (1998). "Computers and language learning: an overview", Language Teaching 31: 57-71.
37. Waslylyk, T. M.; Barbe, W. B.; Lucas, V. H. (1994). "Basic Skills for Effective ommunication. "
38. Williams, D. A. (1999). "Knowing what you don't know: roles for confi-dence measures in automatic speech recognition", Ph.D. thesis, De-partment of Computer Sciences, University of Sheffield, Sheffield, United Kingdom.

الفصل الرَّابع
التَّقييم الآليّ

د. علي علي فهمي

١- تقنيات الأسئلة المقالية وأنواعها.

٢- طرق تقييم المقال (الموضوعات التعبيرية) آليًّا.

٣- تقييم الإجابات القصيرة.

٤- تقييم درجات الكلام.

٥- أنظمة تقييم الرياضيات.

٦- أنظمة الكشف عن السرقات الأدبية.

٧- أنظمة التقييم الآلي ودعم اللغة العربية.

٨- الخلاصة.

التَّقييم الآلي
(Automatic Scoring)

مقدمة

تؤدى زيادة عدد الطلاب والاختبارات إلى أن تصبح عملية تصحيح إجابات الاختبارات بأنواعها وتقييم الطلاب أمرًا مزعجا، ويؤدي التَّقييم الآليّ (Automatic Scoring -AS) إلى اختصار الوقت والجهد، وتوفير تناسق ومتانة التقييم، وتوحيد المقاييس المعيارية لتقييم الطلاب ورصد الدرجات؛ كما أن هذه النظم واسعة المجال بما يكفي لتغطية جميع أنواع إجابات الطالب المكتوبة والمنطوقة.

وتوفر أنظمة التقييم والتصحيح الآلي العديد من المزايا، مثل اتساق وعدالة التقييم، وتقديم تقييمات للامتحانات عالية المخاطر (high-stakes assessments) والتي تتمثل في الحدّ بين القبول والرَّفض «ناجح أو غير ناجح»، «يقبل أو لا يقبل»، «يصلح أو لا يصلح» وهكذا. وتعزز هذه النظم معنى «التوحيد القياسي» عن طريق تطبيق نفس المعايير على جميع الإجابات. بعبارة أخرى فإن التقييم الآلي AES يوفر الفوائد لكافة مهام التقييم بما فيها من مكوناتها الأساسية، والمتَمَثِّلة في الطلاب والقائمين بالتقييم وعملية الاختبار نفسها. وفي الوقت الحالي يتم استخدام نظم التقييم الآلي في تصحيح كثير من اختبارات القبول للجامعات المشهورة مثل اختبارات:

The Test of English as a Foreign Language (TOEFL)، the Graduate Record Examinations (GRE) and the Scholastic Assessment Test (SAT).

تتعامل أنظمة التقييم الآلي AS الحالية مع الطلاب من خلال ثلاث طرق:

الطريقة الأولى هي تقييم إجابات الطلاب المكتوبة، وتشمل تقييم وتصحيح المقالات النصية (ومنها موضوعات التعبير والإنشاء)، تصحيح الإجابات القصيرة.

ويُمَثِّلُ تصحيح المقالات النصية تحدِّيًا أكبر من تصحيح الإجابات القصيرة، حيث يتطلب تقييم أسلوب الطالب في الكتابة style ويحتاج معالجة لغوية عميقة.

الطريقة الثانية هي تقييم إجابات الطلاب المنطوقة، وتنقسم إلى نوعين:

النوع الأول: وفيه يُطلب من الطالب نطق جملة مكتوبة، ويتم تقييم طريقة ودقة النطق. **والنوع الآخَر**: وفيه يطَّلع الطالب على نص مكتوب، ويطلب منه التعبير بصورة منطوقة عن مفهوم النص مستخدما قواعد النطق والقواعد النَّحوية الصحيحة لتكوين الجمل الصحيحة.

يطلق على **النوع الأول** مصطلح «الإجابات المنطوقة بفوضى منخفضة» (low entropy spoken responses) ويطلق على **النوع الآخَر** مصطلح «الإجابات المنطوقة بفوضى مرتفعة» (high entropy spoken responses).

أما **الطريقة الثالثة** للتعامل مع الطلاب فهي تصحيح المسائل الرياضية، والتي تتطلب من الطالب الإجابة بمعادلات ونصوص وأرقام ورسومات بيانية.

ومع أن هدف أنظمة التقييم الآلي AS هو تحقيق علاقة توافق عالية بين الدرجات التي يمنحها الإنسان والدرجات التي تمنحها الآلة، فإنه من المهم أن نعرف أن تقييم درجات إجابات الطلاب تختلف في أسلوبها عند تقييمها بواسطة الآلة عنها عند تقييمها بواسطة الإنسان. فبصفة عامة يوجد أسلوبين لتقييم درجات إجابات الطلاب بطريقة آلية. يعتمد **الأسلوب الأول** على تحديد درجة التطابق التام بين إجابة الطلاب والإجابة الصحيحة النموذجية المحفوظة في النظام. **الأسلوب الآخَر** لا يفترض وجود إجابات نموذجية للمقارنة بها، ولكنه يعتمد على تحليل واستخراج سمات مختلفة من إجابات الطلاب لتحديد نتيجة درجات التقييم الآلي بناءً على التعلم من نتائج التقييم اليدوي التي تمت سابقا على عينة من المقالات.

نهتم في هذا الفصل بنظم التقييم الآلي للأسئلة المقالية والأسئلة التي تتطلب إجابات قصيرة. أما بالنسبة لأنظمة تقويم الأسئلة من نوعية تعدد الخيارات، أو اختيار الإجابة من بين اختيارات الخطأ والصواب، أو التطابق، أو ملء الفراغ؛ فهي سهلة التنفيذ والتطبيق ولن نتعرض لها.

أمكَنَ الإفادة من تقنيات التقييم الآلي للأسئلة المقالية في بناء أنظمة للكشف عن

السرقات الأدبية (Plagiarism detection) حيث تنتشر هذه الظاهرة -للأسف- في الأوساط الأكاديمية، إذ عادةً ما تكون في وثائق المقالات أو التقارير. ومع ذلك، يمكن وجود ظاهرة السرقات الأدبية في أي مجال تقريبا، بما في ذلك الأوراق العلمية، والتصميمات الفنية، وحتى في برامج الحاسب.

يركز هذا الفصل على المنهجيات ونتائج التطبيقات بواسطة كبريات شركات التطوير في مجال التقييم الآلي AS مثل: شركة خدمة الاختبارات التربوية ETS، وشركة معارف تكنولوجيا بيرسون PKT، وشركة فانتيج ليرنينج.

Educational testing Service (ETS), Pearson Knowledge technologies (PKT) and Vantage Learning.

ويشتمل الفصل على عدَّة محاور رئيسية: أنواع الأسئلة المقالية ونظم تقييمها الآلية (AES)، نظم تقييم الإجابة القصيرة، نظم تقييم الكلام (الإجابات المنطوقة)، نظم تقييم أسئلة الرياضيات، أنظمة الكشف عن السرقات الأدبية، وأخيرا.. موقف اللغة العربية من هذا المجال المهمّ.

١- الأسئلة المقالية وأنواعها

تنقسم الأسئلة المقالية إلى أربعة أنواع رئيسية، يعكس كل منها هدفًا تعليميًّا مختلفًا:

١- أسئلة وصفية (Description):

وتتطلب سرد النقاط الرئيسية في الموضوع؛ وغالبا ما تبدأ بالأفعال التالية:

حدِّد، صِف، خطِّط، اشرح، عَدِّد، اذكر، لخص، قدِّم

Define, describe, outline, explain, list, delineate, trace, state, summarize, present

٢- أسئلة مناقشة (Discussion):

وتتطلب مناقشة النقاط الرئيسية في الموضوع؛ وغالبا ما تبدأ بالأفعال التالية:

حلل، استكشف، ناقش، علق، وضح، فسر، استعرض

Analyse, explore, discuss, comment, illustrate, account for, interpret, review, explain, consider, debate, show how and examine

٣- أسئلة تقييم (Evaluation):

وتتطلب إيضاح الخط الفكري أو الحجة خلف الموضوع؛ وغالبا ما تبدأ بالأفعال التالية:

انقد، قيِّم، برِّر، علِّق على

Criticize, evaluate, critically evaluate, justify, comment on, and interpret

٤ - أسئلة مقارنة (Comparison):

وتتطلب مناقشة نقاط التوافق والاختلاف أو نقاط القوة والضعف؛ وغالبا ما تبدأ بالأفعال التالية:

قارن، وضِّح أوجه التَّبايُن، فَرِّق، مَيِّز، ناقِش

Compare, contrast, differentiate, distinguish, debate

٢- طرق تقييم المقال (الموضوعات التعبيرية) آليًّا (AES)

يعرف تقييم المقال آليا بأنه التِّقنية الحاسُوبيَّة التي تقوم بتقييم ووضع درجات للأعمال المكتوبة؛ ويعرف تقييم المقال أيضا بأنه تقدير وتقييم المقالات آليا، ووضع درجات المقالات المكتوبة آليا.

معظم أعمال تقييم المقال آليا تتعامل مع اللغة الإنجليزية، مع قليل من النظم الَّتي صممت لدعم لغات أخرى، مثل: العبرية واليابانية ولغة الملايو «البهاسا».

تقييم المقال (الموضوعات التعبيرية) آليا لا يفترض وجود إجابات نموذجية للمقارنة بها، ولكنها تعتمد على تحليل واستخراج سمات مختلفة من إجابات الطلاب لتحديد نتيجة درجات التقييم الآلي بناءً على التعلم من نتائج التقييم اليدوي التي تمت سابقا على عينة من المقالات.

يمرُّ بناء معظم أنظمة التقييم الآلي للموضوعات التعبيرية عبر نفس خطوات بناء الخوارزمات الإشرافية (supervised algorithms) التي تتطلب مرحلة التعلم. فمراحل بناء نموذج التقييم الآلي AES مقنَّنة، وهي كالتالي:

مرحلة الإعداد، وتشمل تجهيز عينة من المقالات التعبيرية التي تم تصحيحها وتقييمها سابقا بطريقة يدوية مع تحديد واستخراج السمات التي سيتم تدريب النظام عليها من المقال. وتسمى هذه المجموعة: عينة أو فئة التدريب. يتم فحص عينة التدريب هذه من قبل خبراء الحاسب (ويفضل التعاون مع الخبراء في المجال) لتحديد واستخراج مجموعة من خصائص وسمات وأوزان تصف النص المقالي. ففي مرحلة استخراج السمات والتدريب يتم تصحيح وتقييم مجموعة من مئات المقالات التعبيرية بواسطة الخبراء (القائمين بالتقييم) وإعطاء كل مقالة درجة من ١٠ مثلا، وتستخدم هذه السمات والأوزان لإنتاج نموذج رقمي للنص يمكن استخدامه للتنبؤ بدرجة التقييم التي يحصل عليها المقال بواسطة الخبير الإنسان.

مرحلة بناء نموذج برنامج التصحيح والتدريب، حيث يتم استخدام إحدى خوارزمات تعلم الآلة، مثل خوارزم آلة الدعم الموجَّهة (Support Vector Machine) أو الخوارزمات الإحصائية مثل خوارزم بايز (Bayes Algorithm) بغرض تعلم العلاقة الكامنة بين السمات المستخرجة من الموضوع التعبيري وبين درجة التقييم التي تمت يدويا بواسطة الخبير المقَيِّم للمقال. ويتم التحقق من صحة هذا النموذج الرقمي من خلال مقارنة النتائج الَّتي يتم الحصول عليها يدويا من قبل المقَيِّمين الخبراء ودرجة التقييم المستنتجة من هذا النموذج، ويتم تكرار هذه العملية حتى نتأكَّد من تطابق تقييم الحاسب للمقال مع التقييم اليدوي بصورة مُرضية.

وأخيرا مرحلة الاستخدام الفعلي للبرنامج في تصحيح مقالات الطلاب التعبيرية الجديدة آليا. حيث يتم تغذية البرنامج بالسمات المستخلصة من المقال الموضوعي المراد تقييمه آليا.

والآن، كيف يتم تحديد واختيار الخصائص والسمات المعبرة عن نمذجة المقال التعبيري؟ هناك طريقتان أو منهجيتان رئيستان لإنتاج نماذج التقييم الآلي AES، إما باستخدام أساليب «القوة الحاسوبية المحضة» (brute-force) أو باستخدام الوسائل المختلطة (الهجين).

تستخدم المنهجيَّة الأولى تشكيلة واسعة متنوعة من السمات والخصائص اللغوية للنص، والتي ليس لها علاقة مباشرة بكيفية الكتابة الجيدة لمقال، مثل وجود أخطاء من النحو الإملائي، ووجود أخطاء التنقيط، وهكذا.

بينما تكون النماذج القائمة على الأساليب المهجنة ذات علاقة مباشرة مشتقة من الناحية النظرية لمفهوم خصائص الكتابة الجيدة للمقال مثل مدى ارتباط تسلسل الفقرات في النص – وهو ما يطلق عليه مصطلح (Lexical Chaining) – وأن الكلمات المستخدمة هي المناسبة، ومدى استخدام الجمل الموجزة، ومدى استخدام صيغ نائب الفاعل في المقال، واستخدام صيغ الماضي، وهكذا.

وجدير بالذكر أن تحديد واستخراج سمات المقال الموضوعي الملائمة للتقييم هو التحدي الحقيقي لنظم تقييم المقال.

٢,١ – نماذج من أنظمة التقييم الآلي AES

- **نظام «مشروع تصحيح المقال» (Project Essay Grade - PEG)**

وقد تم تطويره بجامعة «كونيتيكت» الأمريكية منذ منتصف الستينيَّات. هو نظام رائد في تقييم المقال آليا AES عبر تاريخ التقييم الآلي. يعتمد هذا النظام على سمات وقياسات تمثل جودة المقالات. هذه القياسات تأخذ في الاعتبار بنية الكتابة مثل متوسط طول الكلمة، ومتوسط طول الجملة بالمقال، إضافة إلى عدد من الوحدات النصية الأخرى.

يستخدم النظام إجراءً إحصائيًّا لإنتاج أوزان ترجيحية لهذه القياسات (باستخدام أُسلوب تحليل الانحدار (Regression Analysis)).

لقد تمَّت إعادة تطوير وتحسين نظام PEG في أواخر التسعينيَّات بإدراج أدوات معالجة اللغات الطبيعية (NLP)، مع الأخذ في الاعتبار التحليل النحوي وعلامات أجزاء الكلام (Parts of Speech tags) للمقال ومدى التزام النص بقواعد النحو.

ويُستخدم نظام PEG في تصحيح برامج الاختبارات تعليمية، مثل امتحان (Scholastic Assessment Test– SAT) الذي يستخدم كاختبار للقبول في الكليات الأمريكية منذ عام ٢٠٠٥، وهو يقيس معلومات وقدرات الطالب في الرياضيات والقراءة النقدية والكتابة.

- **نظام «مقيم المقالة الذَّكيّ» (Intelligent Essay Assessor – IEA)**

لقد تم تطوير المقَيِّم الذَّكِيّ IEA أصلا في جامعة كولورادو الأمريكية في عام ١٩٩٧، ويُسَوَّق حاليا عن طريق شركة (برسون نولدج تكنولوجي PKT). هذا النظام قادر على تقييم المقالات بدرجة ثقة تماثل الخبرة البشرية الماهرة .

يمتلك المقيم الذكيّ IEA العديد من المزايا التي تميزه عن غيره من أنظمة تقييم درجات المقالات، حيث يوفر تقديرا كليا وتغذية عكسية عن الأخطاء الإملائية والنحوية. كما أنه يحتوي في بنيته على كاشفات للمقالات غير العادية، مثل مهارات القيادة العسكرية. وقد جرى استخدامه لرصد الدرجات والتقييم على مدى أكثر من مليون مقالة، تتراوح بين مقالات المدارس المتوسطة ومقالات طُلاَّب كليات الطب، وفي مجالات متنوعة المحتوى.

ميزة هذا النظام أنه يركز على تقييم محتوى المقالة في المقام الأول، لا على التركيب البنائي لها فقط كما هو الحال في النظام السابق. ويقوم مقيم المقالة الذكيّ IEA بتقييم ووضع الدرجات باستخدام أسلوب تحليل الدلالات الكامنة (Latent Semantic Analysis– LSA)، والذي يُمَثِّل طريقة تحليل دلالة النص الَّتي يمكن تعريفها بأنها «نموذج إحصائيّ من استخدام الكلمة الَّتي تسمح بمقارنات التشابه الدلاليّ بين قطع من المعلومات النصية، حيث تنتج مجموعة من المفاهيم المرتبطة بمحتوى النص».

ويفترضُ أسلوب LSA أن الكلمات القريبة في المعنى غالبا ما تكون قريبة من بعضها البعض في داخل النص. ويدمج نظام التقييم الذكيّ IEA طريقة التحليل الدلاليّ LSA جنبا إلى جنب مع قاعدة بيانات معلوماتية تحتوى على مادة الكتب المدرسية وعينة مقالات أو مصادر أخرى غنية في الدلالة لتدريب أجهزة الحاسب.

يؤدي هذا الدمج بين التحليل الدلالي وقاعدة البيانات المعلوماتية إلى تقليل عدد المقالات التي تُستخدم في تدريب النظام، والتي يتم تصحيحها وتقويمها يدويا نظرا لأن رصد الدرجات يتم إنجازه اعتمادا على التحليل الدلاليّ بدلا من بناء النماذج الإحصائية التقليدية للمقال.

- **نظام إنتليمتريك (Intellimetric)**

وهو نظام تم تطويره اعتبارا من ١٩٩٧ بواسطة شركة تكنولوجيا التعلم (Vantage Learning Technology). ويعتبر أول نظام لتقييم المقال آليا AES يستند على علوم الذكاء الاصطناعيّ واللغويات الحاسُوبيَّة؛ فهو يجمع بين أدوات معالجة اللغات الطبيعية (NLP) والتقنيات الإحصائية في رصد درجات المقال، ويمكن الإشارة إليه كمحرك تعلم استوعب «الحكمة الجمعية» (Pooled Wisdom) أو «يستند على العقل» لخبراء التقييم. يستخدم إنتليميترك (Intellimetric) النموذج الذي يحتوى على أمثلة مجموعة من معاملات التنبؤ والأوزان الَّتي تم تعريفها عن طريق استخراج أكثر من ٤٠٠ سمة من أجوبة الطالب، بالإضافة إلى مجموعة سمات تدريبية تتكون من السمات ذات الطبيعة الدلالية والنحوية والخطابية.

يأخذ نظام (IntelliMetric) في الاعتبار خمسة أبعاد أساسية كامنة، وذلك عند تصحيح المقال ورصد الدرجات، وهى: المحتوى، والإبداع، والأسلوب، والميكانيكية، والتنظيم. يستخدم (Intellimetric) شبكات كلمات معجمية مبنية على الإحصاء الدلالي للمقالة. هذا الإحصاء الدلالي يماثل أسلوب التحليل الدلاليّ الكامن (Latent Semantic Analysis-LSA). هذا التحليل الدلاليّ الكامن يمثل خمس فئات عامة من السمات.

الفئة الأولى تهتم بالتركيز على الوحدة والتماسك والتناسق في الغرض والأفكار الرئيسية في المقالة.

الفئة الثانية تختص بمدى اتساع نطاق المحتوى ودعم الأفكار، وتُعنى بمدى الاختيار السليم للمفردات والمفاهيم.

الفئة الثالثة تهتم بمدى تنظيم وهيكلة المقالة من حيث منطق الخطاب، بما في ذلك سيولة الانتقالية والعلاقات بين أجزاء الاستجابة.

الفئة الرابعة وتختص ببنية الجملة والتركيز على تعقيد الجملة والتنوع، مثل: التنوع النحوي في الاستخدام، ومدى التعقيد في الجمل المستخدمة. وأخيرا..

الفئة الخامسة تعكس آليات التحقق من التزام المقال بقواعد اللغة الإنجليزية، مثل: قواعد النحو والإملاء، والحروف الكبيرة، واكتمال الجملة، وعلامات الترقيم، وغيرها.

حاليا يستخدم نظام (Intellimetric) في كثير من المدارس الأمريكية والإنجليزية للصفوف السادسة والسابعة والثامنة، وهو متاح للاستخدام للطلاب عن طريق شبكة الإنترنت.

http: / /www.vantagelearning.com /products /intellimetric / demonstration /demonstration-american-english /

- **نظام التصنيف (E-rater)**

هو نظام تم تطويره بواسطة شركة خدمة الاختبارات التربوية (Educational Testing Service -ETS). وهو معروف جيدا في تنبؤ درجات تقييم المقال ويتفق مع درجات التقييم اليدوية، بالإضافة إلى قدرة النظام على اكتشاف إجابات الطلاب الخارجة عن الموضوع. ويستخدم نظام المصنف E-rater حاليا من أجل:

- تقييم درجات المقالات المرسلة، وتطبيق تعليمات كتابة المقالات في نظام ETS.
- تقييم اختبارات القبول لبرامج الدراسات العليا في الإدارة (Graduate Management Admission Test- ®AWA GMAT). ويقيس هذا الامتحان مهارات الكتابة اللفظية والرياضية، ومهارات الكتابة التحليلية.
- تقديم خدمة تقييم المقال من خلال شبكة الإنترنت. في هذا التطبيق يقوم المحرك برصد درجات المقال ببساطة عن طريق استخراج سمات مستندة على أساس لغويّ من المقال ويستخدم النماذج الإحصائية لربط هذه السمات مع نوعية جودة الكتابة عموما. يتم تقييم نتيجة المقال بدرجة من ١ إلى ٦ حيث ١ هي أدنى درجة و ٦ هي أعلى الدرجات.

يطبق نظام (E-rater) أسلوب الانحدار الخطي المتدرج على عينة من المقالات التدريبية المكتوبة حول نفس الموضوع الذي تم تقييمه بواسطة مجموعة من المتخصصين من أجل استخراج أكثر من ٥٠ سمة لغوية للمقال، والتي يمكن أن تكون عونا كبيرا في التنبؤ بتقييمات المقالات المماثلة في نفس الموضوع.

يطبق نظام (E-rater) المحدَّث مجموعة من السمات التي يستخلصها من المقال موزعة إلى خمسة مجالات من التحليل؛ الأول سمات خاصة بالأخطاء النحوية وأخطاء الاستخدام، والأسلوب Style. الثاني هو تنظيم المقال. الثالث هو تعقيد المعجم. الرابع هو مدى الاستخدام الصحيح للمفردات، وأخيرا طول المقال. يتضمن نظام التصنيف (E-rater) سمات تصنيف أخرى تتعلق بالمفردات ومدى ملائمة المحتوى، والتنظيم، والتطوير.

- **نظام التصنيف (C-rater™)**

تم تطوير هذا النظام من قبل شركة خدمة الاختبارات التربوية (Educational Testing Service -ETS) أيضا، وهو معروف جيدا بالتقييم ذي الدِّقَّة العالية للمقالات المكتوبة، وتم التحقق منه على مقالات متعددة من برامج الاختبار في العديد من مجالات المحتوى المختلفة، بما في ذلك العلوم، والقراءة والفهم والتاريخ.

تستخدم تكنولوجيا نظام التصنيف (C-rater) «منهجية حقيبة الكلمات» (Bag of words approach) والتي تستخدم فيها المعالجة الطبيعية العميقة للغة لتقييم ما إذا كانت إجابة الطالب تحتوى على النص الذي يمكن اعتباره صياغة أخرى مماثلة للمفاهيم الواردة في شرح الموضوع (item rubric). يختلف هذا المنهج عن الأساليب الأخرى لتحليل إجابات الطلاب (مثل تحليل الدلالة الكامنة LSA).

لإجراء عملية التقويم يقوم نظام (C-rater) بإجراء سلسلة من خطوات معالجة اللغات الطبيعية NLP ومنها :

- تصحيح الأخطاء الإملائية الطلاب.
- تحديد بنية كل جملة نحوية.
- حل مرجع الضمير.
- تحليل صياغة إجابات الطلاب.

الميزة الرئيسية لمحركات (C-rater) عن باقي محركات AES الأخرى هي التحليل اللغوي العميق لإجابات الطلاب، وهو ما يضمن أن عملية التقييم لن تنخدع بالإجابات الَّتي تستخدم الكلمات الصحيحة في سياق خاطئ.

٢,٢- نتائج تطبيقات أنظمة التقييم الآلي AES

تقاس نتائج التقييم الآلي بمدى تطابقها مع التقييم اليدوي، مع الأخذ في الاعتبار أنه نادرا يندر تطابق نتائج التقييم لشخصين مختلفين.

ويمثل الجدول التالي نتائج التطبيقات المحققة من حيث الاختبار، وحجم عينة مقالة التقييم، وارتباط الإنسان بالإنسان وارتباط الإنسان والحاسب.

النِّظام	الاختبار	حجم العينة	ارتباط الإنسان والإنسان	ارتباط الإنسان والحاسب
PEG (1997)	GRE	٤٩٧	٠,٧٥	٠,٧٥ – ٠,٧٤
PEG (2002)	English placement test	٣٨٦	٠,٧١	٠,٨٣
IntelliMetric (2001)	k-12 norm-referenced test	١٠٢	٠,٨٤	٠,٨٢
IEA (1997)	GMAT	١٨٨	٠,٨٣	٠,٨٠
IEA (1999)	GMAT	١٣٦٣	٠,٨٧ – ٠,٨٦	٠,٨٦
IEA (2011)	High School Writing	٦٣٥	٠,٩١	٠,٩١
e-rater (1998)	GMAT	١٠٠٠ – ٥٠٠	٠,٨٩ – ٠,٨٢	٠,٨٧ – ٠,٧٩
e-rater (2006)	GMAT – TOEFL	٧٥٧٥	٠,٩٣	٠,٩٣
e-rater (2011)	GRE- TOEFL	> ٥٠٠٠	٠,٩٥	٠,٩٧

الجدول ٤-١: نتائج تطبيقات أنظمة التقييم الآلي AES

يتضح من الجدول السابق أن أنظمة التقييم الآلي للمقالات التعبيرية المكتوبة باللغة الإنجليزية قد بلغت درجة عالية من النضج؛ فنتائجها تختلف مع نتائج التقييم اليدوية بنفس القدر (تقريبا) الذي تختلف فيه نتائج التقييم من شخص إلى شخص آخر.

٣- تقييم الإجابات القصيرة.

إنَّ نظم تصنيف درجات الإجابة القصيرة سهلة التنفيذ حيث أنها تهدف إلى تقييم محتوى المعارف والمهارات الطالب، في مقابل نظام درجات المقال الَّتي تقوم بتقييم قدرة الطالب على الكتابة وتتطلب إمكانيات متطورة لتحليل النص وفهمه. نظم تقييم الإجابات القصيرة تتطلب أن تكون إجابة الطالب قصيرة على أن تبين مدى استيعابه للمفاهيم الرئيسية في مجال معين. ويقوم نظام التقييم الآلي بمقارنة إجابة الطالب مع واحد أو أكثر من الأجوبة الصحيحة المخزنة لديه. في الماضي القريب كانت معظم أنظمة التقويم تتطلب مجهودًا إضافيًّا من المعلم حيث كان يتطلب منه إعداد أدلة الإجابات النموذجية بطريقة يدوية أو أن يوفر مُدَوَّنة لُغَوِيَّة مُعَنونة (Annotated Corpus) لتحديد أنماط الإجابة بطريقة نصف آلية.

تعتبر منهجية قياس تشابه النص (Text similarity approach) هي أساس عمل نظم تقييم الإجابات القصيرة. ويوجد عدد كبير من خوارزمات التشابه بعضها يأخذ في الاعتبار التحليل اللغوي العميق لكل من إجابة الطالب وإجابة المدرس والبعض الأخر يأخذ في الاعتبار العبارات القصيرة المشركة بين الإجابتين.

يلعب التشابه الدلالي بين كلمتين دورا كبيرا في الوصول إلى التشابه الدلالي بين جملتين (غالبا ما يتم ذلك باستخدام تقنية المعلومات المتبادلة بين كلمات الجملتين pointwise mutual information). ولكن كيف يمكن حساب التشابه الدلالي بين كلمتين (مثل الكلمتين شجرة، نخلة أو كتاب وكراسة)؟

توجد طرق كثيرة لذلك نذكر منها الطرق التالية:

- Leacok & hodorow.
- Lesk.
- Wu& Palmer.
- Resnik.
- Lin.
- Jiang & Conrath.
- Hirst & St-Onge.

- Corpus based Similarity combined with Explicit Semantic Analysis (ESA).
- Corpus based Similarity combined with Latent Semantic Analysis (LSA).

فعلى سبيل المثال يقيس Lesk التشابه بين كلمتين عن طريق حساب نسبة التداخل بين التعريف المناظر لكل كلمة على النحو المنصوص عليه من قبل القواميس الشهيرة.

وأيضا يحدد (Wu& Palmer) التشابه بين كلمتين عن طريق قياس مدى التباعد بين موقع الكلمتين في شبكة الكلمات المعجمية للغة الإنجليزية (WordNet).

وفيما يلي نماذج من أنظمة درجات الإجابات القصيرة

٣,١- نظام أكسفورد (UCLES)

يستخدم هذا النظام مجموعة من الكلمات والمترادفات ونوافذ البحث والمطابقة مع أنماط الإجابة النموذجية، ويتم تعلم النظام باستخدام عدة أساليب لتعلم الآلة مثل أسلوب شجرات القرار (Decision trees)، وتقنيات بايز (Bayesian Learning)، وأساليب البرمجة المنطقية الاستقرائية (Inductive Logic Programming).

لتقييم مدى جودة نظام أكسفورد فقد تم تجربته على تقييم إجابات تسع أسئلة من مقرر الأحياء في شهادة الثانوية العامة الأمريكية (General Certificate of Secondary Education -GCSE). تم تجميع ٢٦٠ إجابة تجريبية لكل سؤال من الأسئلة التسعة. وتراوحت علامات الدرجات لهذه الأسئلة من ١ إلى ٤.

ومن بين ٢٦٠ إجابة لكل سؤال تمَّ اختيار ٢٠٠ إجابة والدرجة المصححة لكل منها وذلك لتدريب نظام أكسفورد على كيفية الإجابة (من خلال بناء نموذج لتعلم الآلة) واستخدمت ٦٠ إجابة متبقية في اختبار النظام. أوضحت النتائج تطابق نتيجة التصحيح بواسطة نظام أكسفورد مع نتائج التصحيح بواسطة مدرسي المقرر بنسبة ٨٤٪.

٣,٢- نظام التقييم (C-rater)

بالإضافة إلى كون (C-rater) يستخدم في تقييم الأسئلة المقالية، إلا أنه يستخدم أيضا في تصحيح الأجوبة القصيرة. وهو نظام وضع درجات آلي، يستخدم التحليل الصرفي والمترادفات، وهيكل الإسناد والوسيط ومرجع الضمير لتقييم الإجابات القصيرة المعتمدة على المحتوى.

تم تقييم كفاءة نظام التصنيف (C-rater) من خلال برنامجين للتقييم واسعة النطاق. كان البرنامج الأول هو مشروع التقييم الوطني للتقدم التعليمي بالولايات المتحدة في الرياضيات (National Assessment of Educational Progress- NAEP). وقد استُخدم نظام التصنيف (C-rater) لتقييم إجابات الطلاب الخاصة بتفسير الحلول الخاصة لبعض مسائل الرياضيات. حيث كان متوسط طول إجابات الطلاب حوالي ١,٢ جملة أو عدد ١٥ كلمة.

البرنامج الآخَر هو برنامج إدارة وتصحيح مقرر الإنجليزية في جامعة إنديانا الأمريكية من خلال شبكة الإنترنت. في هذه الحالة، كان مطلوبا من برنامج (C-rater) تقييم سبعة أسئلة لفهم المقروءات، حيث الإجابات على هذه الأسئلة أكثر انفتاحا من إجابات الأسئلة المتعلقة بفهم الحلول الرياضية NAEP.

تم اختيار وتقييم استجابات الطلاب بين ٢٤٥ و ٢٥٠ عشوائيا من قبل اثنين من المصححين البشر وبواسطة نظام التصنيف C، وكان معدل اتفاق النظام مع المصحح البشري الأول ٨٤,٤٪ بينما كان معدل الاتفاق بين النظام والمصحح البشري الثاني ٨٣,٦٪. وكان معدل الاتفاق المتوسط بين الاثنين من المصححين البشر ٩٠,٨٪.

وهذا يعني أن أداء نظام التصنيف C كان مشجعا في حالة تقييم الأسئلة الموضوعة من قبل منظومة التقييم الوطني للتقدم التعليمي بالولايات المتحدة NAEP.

٣,٣- نظام تسجيل الدرجات آليا (Automark)

وهو نظام برمجيات يستخدم تقنيات معالجة اللغات الطبيعية NLP لإنجاز وضع الدرجات المحَوسَبة على إجابة نص حر على أسئلة مفتوحة، ويستخدم تقنيات استخراج المعلومات (Information Extraction) لاستخلاص المفهوم الكامن أو المعنى وراء

النص الحر. يستند وضع الدرجات فيه أساسا إلى تحليل المحتوى مع الوضع في الاعتبار بعض سمات الأسلوب الَّتي ينبغي النظر فيها. تمر عملية وضع الدرجات من خلال أربع مراحل.

أولا: تتم المعالجة المسبقة لإجابة الطالب حتى تكون موحدة من حيث الإملاء وعلامات الترقيم، والتأكد من أن النظام يتسامح مع أخطاء الإملاء والكتابة وبناء الجملة.

ثانيا: يقوم محلل الجملة بالتعرف على المكونات النحوية الرئيسية للنص وكيفية ارتباطها.

ثالثا: ويقوم جزء برمجية تطابق الأنماط بالبحث عن التطابقات بين قوالب نظام وضع الدرجات والمكونات النحوية المكونة لنص الطالب.

رابعا: وفي نهاية المطاف تقوم وحدة «**التعليق على إجابة الطالب**» بمعالجة نتيجة تطابق الإجابة مع النمط المخزن، ويكون **التعليق على إجابة الطالب** في صورة الدرجة الَّتي يحصل عليها، ومن الممكن أن تكون أكثر تحديدا.

تم اختبار التقييم الآلي (Automark) لتقييم المناهج الوطنية الأمريكية للعلوم لتلامذة في عمر الإحدى عشرة سنة (National Curriculum Assessment of Science for eleven years old pupils). وكان شكل إجابات الطلاب : كلمة واحدة، قيمة واحدة، وصف الجملة التفسيرية القصيرة، أو وصف النمط الموجود في مجموعة من البيانات. وتراوحت علاقة الارتباط المتحققة بين ٩٣٪ و ٩٦٪ مقارنة بالتصحيح اليدوي.

٤- تقييم درجات الكلام

تقييم درجات الكلام آليا يشبه إلى حد كبير تقييم المقال المكتوب آليا.

أولا، يتم استخراج سمات اللغة ذات الصلة، ومن ثم يتم استخدام نموذج لحساب الدرجات على أساس مزيج من هذه السمات. يختلف التقييم الآلي للمقال المكتوب عن تقييم الكلام المنطوق في نقطتين رئيسيتين النقطة الأولى: أن تقييم الكلام المنطوق

يتطلب برمجة إضافية لتحويل الكلام إلى كتابة. النقطة الثانية: عادة ما يكون اختبارات الكلام لغير ناطقي اللغة الأصلية (Non native speakers).

تصنف مهام رصد درجات الكلام في فئتين أساسيتين : مهام فوضى منخفضة ومهام فوضى قُصوى.

تقوم مهام الفوضى المنخفضة برصد درجات الاستجابات الَّتي يمكن التنبؤ بها إلى حد كبير مثل:

- القراءة الشفوية من فقرات مكتوبة.
- طلب تكرار جملة منطوقة مسجلة مسبقا على الحاسب.
- طلب إجابة منطوقة لأسئلة محددة الإجابة (Factual Questions).
- طلب وصف صورة بسيطة.

في المقابل فإن مهام الفوضى القُصوى تقوم بالتعامل مع الكلام المتجدد - مقصودًا كانَ أم عفوِيًّا.

وفيما يلي نماذج من أنظمة تقييم الكلام:

٤ ,١ - محرك أو آلة مصنف الكلام (SpeechRater ETS Engine)

يعتبر مُحَرِّك (SpeechRater) أحد تطبيقات رصد درجات الاستجابة؛ حيث يستخدم لتقدير درجات الاستجابات العفوية، بما في ذلك مجال الاستجابات الممكنة مفتوحة النهاية على النقيض من الإجابات المقيدة.

وقد تم تقييم درجات المتقدمين لاختبار الاستعداد للحصول على اختبار تويفل عن طريق تقدير الدرجات باستخدام آلة مصنف الكلام كجزء من اختبار ممارسة تويفل TOEFL من خلال الإنترنت منذ عام ٢٠٠٦. تركز مسابقات تقييم أنظمة التعرف على الكلام وتقييمه على الجوانب ذات المستوى المنخفض من إنتاج الكلام مثل النطق (pronunciation) باستخدام مهام مقيدة من أجل زيادة الموثوقية في النظام. على النقيض من ذلك فإن محرك مصنف الكلام (The SpeechRater)، يعتمد على مفهوم واسع لبناء إجادة الحديث بالإنجليزية، ويشمل جوانب التوصيل الجيد للكلام (مثل

الطلاقة في الحديث ودقة النطق)، وتسهيلات قواعد اللغة والقدرات رفيعة المستوى التي تتعلق بالتماسك الموضعيّ وتطور الأفكار.

يعالج محرك مصنف الكلام (SpeechRater engine) كل استجابة مع نظام التعرف الآلي على الكلام المكيف خصوصا للاستخدام مع الإنجليزية العامية. واستنادا إلى مُخرج هذا النظام، يتم استخدام معالجة اللغات الطبيعية لحساب مجموعة من السمات التي تميز «صورة» من الكلام بناء على عدد من الأبعاد اللغوية، بما في ذلك الطلاقة، واستخدام المفردات، والنطق واللحن في الكلام. تستخدم هذه من أجل تعيين درجة تقييم نهائيّ لإجابة الطالب. بينما يتم تشييد بنية هذا النموذج من التقييم من قبل خبراء المحتوى، فإنه أيضا يتم التدريب على قاعدة بيانات من نتائج تقييم إجابات سابقة بواسطة خبراء في المجال وذلك من أجل ضمان أن يحقق محرك مصنف الكلام محاكاة عالية لتقييم الإنسان بقدر الإمكان. علاوة على ذلك، إذا اكتشف النظام عدم قدرته علي تقييم إجابة الطلاب المنطوقة نتيجة مُشكلات جودة الصوت أو أيَّة مُشكلات أخرى، فإن محرك تصنيف الكلام يمكنه وضعها جانبا للمعالجة الخاصة.

وتسعى شركة (Educational Testing Service –ETS) صاحبة نظام SpeechRater إلى تطوير العديد من خصائص المعالجة الطبيعية للغة (NLP) لتمثيل التعبيرات النحوية واكتشاف ملامح هيكل استجابة الردود المنطوقة. كما تسعى إلى زيادة قدرة النظام للاستخدام على نطاق واسع من المستخدمة في عمليات تقييم كفاءة الكلام باللغة الإنجليزية، بما في ذلك مجموعة من الخيارات المقيدة جداً (مثل قراءة فقرات مكتوبة بصوت عال)، مرورا بالبنود الأقل تقييدا (مثل مهام تلخيص وقراءة نص مكتوب)، إلى الخيارات المفتوحة بشكل كامل.

٤ , ٢– محرك أو آلة مصنف الكلام فرسانت (PKT Versant)

تطبيق فرسانت (PKT Versant) هو اختبار آلي للغة المنطوقة، والذي يمكن تنفيذه بسهولة عبر الهاتف أو جهاز الحاسب من قبل مجموعات كبيرة من المرشحين. ويتم رصد درجات الاختبارات آليا في غضون دقائق تلقائيا، ويتمكن هذا التطبيق من تنفيذ كل من: تقديم نتيجة الرصد عموما، بالإضافة إلى رصد درجات المهارة الأعلى. وقد ساعدت اختبارات التطبيق كلا من الشركات والوكالات الحكومية والجامعات

والمدارس في القياس الدقيق والسريع لمهارات التحدث بالإنجليزية، أو بالإسبانية، أو بالعربية في أكثر من ١٠٠ دولة حول أنحاء العالم من أجل أغراض الاختبارات والتدريب.

يقوم نظام اختبار فرسانت (Versant) آليا بتقييم ردود الاستجابات للعديد من المهام المختلفة. تتضمن اختبارات نظام فرسانت: القراءة بصوت عال، وتكرار الجمل، وبناء الجمل، وإعطاء الأجوبة على أسئلة قصيرة، رواية قصص قصيرة، واختيار الاستجابة، والمحادثة، وطريقة القراءة والفهم. في اختبار الكتابة لهذا التطبيق فإن بند المهام يتضمن: الكتابة، وإكمال الجمل، والإملاء، وإعادة البناء، وكتابة رسائل البريد الإلكتروني. بالنسبة لبعض المهام، مثل القراءة والتكرار فهناك سلسلة واحدة من تتابع الكلمات هي بالضبط الصحيحة المتوقعة لكل استجابة. في مهام أخرى، يمكن أن تكون العناصر متعددة الأجوبة الصحيحة. تمر كافة عناصر الاختبار بإجراء اختبار مسبق عالي الكثافة على عينات مختلفة متنوعة من الناطقين وغير الناطقين باللغة في طائفة واسعة النطاق من مجال مستوى القدرة.

٤, ٣- محرك أو آلة مصنف الكلام إديو سبيك (EduSpeak)

نظام إديو سبيك (EduSpeak) من SRI الدولية نظام مجموعة أدوات تطوير برمجيات تمكن مطورو البرمجيات من تعليم اللغة تفاعليا باستخدام أحدث نظم تكنولوجيا التعرف على الكلام والنطق وتسجيل الدرجات.

يسمح رصد درجات النطق الآلي للحاسب بتقديم ملاحظات التغذية العكسية (Feedback) على الجودة النوعية الشاملة للنطق للإشارة إلى مُشكلات التوليد المحددة. كما يسمحُ بمعاينة المنهج في رصد درجات النطق، حيث إن الهدف هو تقدير درجة التقييم لنوعية نطق فقرة أو جملة يسعى الخبير البشري إلى تعيينها. وتدعم أدوات إديوسبيك (EduSpeak) وظيفة استشعار خطأ النطق على مستوى الهاتف باستخدام قواعد البيانات من الكلام وتقييمات الإنسان على مستوى الجملة، حيثُ إنَّ بعض المقاطع عبر الهاتف تفتقر إلى جودة النطق، ومن الممكن تزويد الطالب بالتغذية العكسية وملاحظات حول عن أخطاء نطق محددة.

جرى تقييم قدرة النظام على الكشف عن أخطاء النطق (mispronunciation) في قاعدة بيانات صوتية من ١٣٠ ألف هاتف لجمل حديث متصل قالها ٢٠٦ شخص من غير الناطقين باللغة الأم وتم تحويلها إلى نصوص مكتوبة. أظهرت النتائج أن نسبة الخطأ أعلى قليلا من الخطإ البشري.

٥- أنظمة تقييم الرياضيات

في مجال الرياضيات، فإن أداء نظم رصد الدرجات آليا قوية عندما يتم تقييد شكل الاستجابة. تتعامل نظم تقييم الرياضيات مع بنود الرياضيات التي تتضمن المعادلات أو التعبيرات الرياضية، والأشكال الهندسية ثنائية الأبعاد والخطوط المتصلة وغير المتصلة أو الخط المنحنى والرسوم البيانية والأشرطة، ومُدخلات الأرقام.

يشهد المجال حاليا ارتفاعًا في جودة هذه الأنظمة، ومن المتوقع أن تنجز هذه النظم مهامها بدقة عالية دون الحاجة لمراجعة المصحح البشري.

وفيما يلي نموذج لإحدى الأسئلة الاختبارات وتتطلب إجاباتها رسومًا بيانية تقوم نظم التقييم الآلية بتصحيحها ومنحها درجة تقييم:

«عائلة تسافر بسرعة ثابتة خلال رحلة الطريق. بعد ٣ ساعات من السير توقف لمدة ساعتين لتناول الطعام والراحة. ثم استأنفت السفر لمدة ٤ ساعات أخرى بنفس السرعة. ارسم رسم بياني تمثل به هذا الوضع».

وهذا نموذج آخر لأسئلة تتطلب صياغة الإجابة في صورة تعبير ومعادلات رياضية:

«في يوم واحد، باع أحد المتاجر عدد ٣٠٠ قميص بتخفيض قيمته ٢٥٪ من السعر العادي وهو س للقميص الواحد.

عبر بصورة رياضية عن إجمالي المبلغ الذي حصل عليه المتجر في ذلك اليوم».

وفيما يلي نماذج من أنظمة تقييم الكلام:

٥,١- نظم تقييم الرياضيات (m-rater)

نظام محرك مصنف درجات الرياضيات من شركة (ETS's m-rater) وهو محرك يستعمل في رصد الدرجات للاستجابات الرياضية مفتوحة النهاية، مثل تلك التي تأخذ شكل تعبيرات أو معادلات رياضية، أو رسوم بيانية. منذ أواخر تسعينيَّات القرن العشرين ١٩٩٠.

٥,٢- نظم تقييم الرياضيات (MathQuery)

محرك ماث كويرى (MathQuery) من شركة بيرسون (Pearson) يعمل في بيئة الإنترنت ويقوم بتقييم مهارات التفكير الحرجة في الرياضيات حيث يتعامل مع مسائل العالم الحقيقي التي يمكن حلها بأكثر من طريقة واحدة والتي يمكن أن تكون لها حلول صحيحة متعددة وليس بالضرورة أن تكون هذه الحلول متكافئة.

يقوم المحرك (MathQuery) بتحليل سلسلة الخطوات أو المسار إلى الحل. وبالنسبة للتعبيرات الرياضية، يقدم المحرك محرر لكتابة المعادلات يمكن تخصيصه لمختلف مستويات الدراسة وموضوعات المحتوى ومزود برموز الجبر والرموز اللازمة لحساب التفاضل والتكامل وغيرها من المقررات الرياضية.

٦- أنظمة الكشف عن السرقات الأدبية (Plagiarism Detection Systems)

تستخدم هذه الأنظمة في الكشف عن السرقات الأدبية التي تتم في المقالات العلمية والأدبية. وتعتمد هذه الأنظمة في تقنياتها على التشابه بين نص ما وبين النص الأصلي الذي سبق نشره في تاريخ سابق. يتم فحص التشابه على نطاقين: النطاق الأول على مستوى النص ككل مثل أسلوب الكتابة (Stylometry)، والنطاق الثاني على مستوى الفقرات وهو ما يطلق عليه المستوى المحلي.

وتقنية بصمة النص (Fingerprint) على مستوى النص ككل هي الأكثر انتشارا. تقوم أنظمة الكشف عن السرقات الأدبية ببناء قواعد بيانات لبصمة النص لعدد ضخم من المقالات المنشورة. وفي حالة فحص مقالة أو نص جديد يتم مقارنة بصمتها مع قاعدة البيانات. في حالة الكشف الإيجابي يتم الفحص التفصيلي بين هذا النص وبين

النصوص المكتشفة، فإذا كانت نسبة التشابه أعلى من قيمة معينة فيعتبر ذلك مؤشرًا قويًّا على وجود سرقة أدبية.

و السرقات الأدبية تنقسم إلى الأنواع التالية:

١- نسخ ولصق (Copy and Paste Plagiarism).

٢- الاحتيال المتنكر (Disguised Plagiarism).

٣- الاحتيال عن طريق إعادة الصياغة (Paraphrase Plagiarism).

٤- الاحتيال عن طريق الترجمة (Translation Plagiarism).

٥- سرقة الأفكار (Idea Plagiarism).

تتجه البحوث حاليا إلى الكشف عن السرقات الأدبية عن طريق الترجمة، وهو ما يطلق عليه (Cross Lingual Plagiarism Detection –CLPD).

وفيما يلي قائمة ببعض الأنظمة المستخدمة للكشف عن السرقات الأدبية:

أنظمة تجارية	أنظمة متاحة للجمهور
Attributor	Chimpsky
Copyscape	CitePlag
Iparadigms: Ithenticate, Turnitin	CopyTracker
Plagiarismdetect	eTBLAST
PlagScan	Plagium
Urkund	SeeSources
Veriguide	The Plagiarism Checker

٧- أنظمة التقييم الآلي ودعم اللغة العربية

توجد جهود بحثية قليلة جدا في هذا المجال رغم أهميته التعليمية وترجع صعوبة تنفيذ أنظمة التقييم الآلي التي تدعم اللغة العربية إلى أنها تتطلب معالجات لغوية عميقة وهي غير متوفرة حتى يومنا هذا بصورة مرضية.

ولكن هذا لا يمنع من البدء في بناء أنظمة تقييم الأجوبة القصيرة وهي لا تتطلب العمق التحليلي اللغوي كما هو الحال بالنسبة للأسئلة المقالية. ولكنها تتطلب وجود ما هو مماثل لشبكة الكلمات للغة الإنجليزية (WordNet) وشبكة الكلمات الدلالية للغة الإنجليزية (SentiWordNet).

٨- الخلاصة

تتنوع مجالات التَّصحيح الآليّ للامتحانات. وقد تم تقديم مختلف النظم الآلية في هذا الفصل في جميع مجالات التقييم ورصد الدرجات. إنَّ دقة النظم هي علاقة ارتباط بين رصد الدرجات بشريا ورصد الدرجات بواسطة النظام. أصبحت نظم رصد الدرجات آليا واقعا، وطالما هناك فرق بين رصد درجات البشريّ ورصد الدرجات الآلي فإن موضوع الدقة هي نقطة جيدة للبحث.

تم تعريب أدوات ونظم وتطبيقات وحزم البرمجيات الجاهزة من خلال تزويدها بالحروف المطبعية (Fonts) الخاصة باللغة العربية والقدرة على تداول الحروف العربية جنبا إلى جنب مع الحروف اللاتينية مع الأخذ في الاعتبار خصائص كتابة اللغة العربية من حيث (أ) اتجاه الكتابة (من اليمين إلى اليسار)، (ب) ومن تغير شكل الحروف طبقا لموقعه في الكلمة، (ج) ومن حيث ترتيب شفرة الحروف (يأتي حرف السين قبل حرف الشين مثلا).

في نظام تشغيل الحاسب - مثل: نظام ويندوز ميكروسوفت - يُوَفِّر نظام التشغيل الخصائص السابقة لمعظم البرمجيات والتطبيقات الَّتي تعمل تحت مظلته، إلا أن التطبيقات الحديثة، مثل : إدارة المعرفة، آلات البحث الذكية، تحويل النصوص المكتوبة إلى نصوص منطوقة، وتصحيح درجات الطلاب.. لا تكتفي بالتعامل مع النصوص على مستوى الحرف والكلمة لكنها تتعامل مع الجملة شكلا ومعنى.

لا يصلح مع التطبيقات الحديثة أسلوب التعريب على مستوى الحرف للتعامل مع اللغة الأم وهى اللغة العربية، حيثُ يستدعي تعريب هذه التطبيقات أخذ شكل ومعنى الكلمة والجملة العربية في الاعتبار، كما أن الاكتفاء بالمعالجة على مستوى الحرف غير كاف في التطبيقات الذكية والمستقبلية، وهناك قصور شديد في تعريب هذه التطبيقات يجب تداركه في المستقبل القريب بقدر الجهد والاستطاعة.

ببليوجرافيا مرجعيَّة

1. Bennett, R. E. (2011). Automated Scoring of Constructed-Response Literacy and Mathematics Items, White Paper, Publisher: Arabella Philanthropic Advisors.
2. Bernstein, J.; Suzuki, M.; Cheng, J.; Pado, U. (2009). Evaluating diglossic aspects of an automated test of spoken modern standard Arabic. ISCA International Workshop on Speech and Language Technology in Education (SLaTE 2009).
3. Bernstein, J.; van Moere, A.; Cheng, J. (2010). Validating automated speaking tests. Language Testing.
4. Blokdyk, G. (2018). Text Mining Complete Self-Assessment Guide. Emereo Pty Limited.
5. Burstein, J.; Chodorow, M. (2010). Progress and new directions in technology for automated essay evaluation. In R. Kaplan (Ed.). The Oxford Handbook of Applied Linguistics (2nd Ed., pp. 487–497). New York: Oxford University Press.
6. Chen, L.; Tetreault, J.; Xi, X. (2010). Towards Using Structural Events to Assess Non-Native Speech, NAACL-HLT 2010: Proceedings of the 5th Workshop on Building Educational Applications (BEA-5) Association for Computational Linguistics.
7. Dikli, S. (2006). An Overview of Automated Scoring of Essays, The Journal of Technology, Learning, and Assessment, Volume 5, Number 1.
8. Feldman, R.; Sanger, J. (2006): The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press.
9. Franco, H.; Bratt, H.; Rossier, R.; Gadde, V. R.; Shriberg, E.; Abrash, V.; Precoda, K. (2010). “EduSpeak: a speech recognition and pronunciation scoring toolkit for computer-aided language learning applications,” Language Testing, vol. 27, no. 3, p. 401.

10. Gomaa, W.; Fahmy, A. (2013): A Survey of Text Similarity Approaches. International Journal of Computer Applications 68(13):13-18.
11. Gomma, W.; Fahmy, A. (2011). Tapping Into the Power of Automated Essay Scoring, 11th Conference on Language Engineering, ESOLE.
12. Gomma, W.; Fahmy, A. (2014). Arabic Short Answer Scoring with Ef-fective Feedback for Students. International Journal of Computer Ap-plications 86 (2):13-18.
13. Hui, E. (2018). Learn R for Applied Statistics: With Data Visualizations, Regressions, and Statistics. Apress.
14. Jo, T. (2018). Text Mining: Concepts, Implementation, and Big Data Challenge. Springer.
15. Kim, J. (2019). Genome Data Analysis. Springer Singapore.
16. Konchady, M. (2006). Text Mining Application Programming (Pro-gramming Series). Charles River Media.
17. Mohler, M.; Mihalcea, R. (2009). Text-to-text Semantic Similarity for Automatic Short Answer Grading, EACL '09 Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics
18. Shermis, M. D.; Burstein, J.; Higgins, D.; Zechner, K. (2009). Automated Essay Scoring: Writing Assessment and Instruction (2010 Elsevier Ltd).
19. Stein, B.; Sven, M.; Potthast, M. (2007). "Strategies for Retrieving Plagiarized Documents", Proceedings 30th Annual International ACM SIGIR Conference, ACM.
20. Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. Language Teaching Research, 10(2), 157–180.
21. Zechner, K.; Higgins, D.; Xi, X.; Williamson, D. (2009). Automatic Scoring of Non-Native Spontaneous Speech in Tests of Spoken

English, Speech Communication, Educational Testing Service, Automated Scoring and NLP, Rosedale Road, MS 11-R, Princeton, NJ 08541, USA 2009 , Vol. 51, No. 10, pp. 883–895.

22. Zizka, J.; Darena, F.; Svoboda, A. (2019). Text Mining with Machine Learning. Taylor & Francis Group.

الباحثون



الدكتور/ مُحسِن عبد الرَّازق علي رشوان



يشغل منصب أستاذٍ بقسم الإلكترونيات والاتصالات الكهربائية في كُلِّيَّة الهندسة – جامعة القاهرة. تخرَّجَ عام ١٩٧٧ وكان الأول على دفعته، وحصلَ على ثلاثة ماجستيرات، ثم على الدُّكتوراه من جامعة كوين بكندا؛ أشرف على أكثر من مائة رسالة ماجستير ودكتوراه. يدير الشَّركة الهندسيَّة لتطوير النُّظُم الرَّقمِيَّة RDI المتخَصِّصة في مجال تقنيات اللُّغة العربيَّة.

الدكتور/ المُعتزّ بالله السَّعيد طه



أستاذ الدِّراسات اللُّغويَّة الـمُساعد بجامعة القاهرة، وأستاذ اللِّسانيَّات الحاسوبيَّة المُشارك بمعهد الدَّوحة للدِّراسات العُليا، ومُنسِّق وَحدة الموارد الـمُعجميَّة بمشروع مُعجم الدَّوحة. نَشَرَ نحوَ ثلاثينَ ورقة علميَّةٍ، بالإضافة إلى عددٍ من الكتب في الـمُعجَمِيَّة العَرَبيَّة والدِّراسات اللُّغَوِيَّة الـمُعاصِرة، وأسهمَ في أكثر من عشرة مشرُوعاتٍ بحثيَّةٍ دوليَّةٍ في ميادين مُعالَجَة اللُّغات الطَّبيعيَّة. حصلَ على عددٍ من الجوائز في مَيدان تخصُّصه، منها: جائزة (ألكسو ALECSO) للإبداع والابتكار في «المَعلُوماتيَّة والـمُعالَجَة الآليَّة للُّغة العربيَّة»، وجائزة راشد بن حُمَيد للعُلُوم والثَّقافة.

الدكتور/ مُحمَّد عَطِيَّة مُحمَّد العَرَبيّ



حصل من جامعة القاهرة على بكالوريوس هندسة الاتصالات الكهربائية والإلكترونيات عام ١٩٩٥م، وعلى الماجستير في هندسة الحاسبات عام ٢٠٠٠م، ثم على درجة الدكتوراه في هندسة الاتصالات الكهربائية والإلكترونيات عام ٢٠٠٥م. عَمِلَ بالشَّركة الهندسيَّة لتطوير النُّظُم الرَّقمِيَّة RDI منذ يوليو ١٩٩٥م إلى نهاية ٢٠١٠م، وبينَ عامَيْ ٢٠٠٧م و ٢٠١٠م أستاذًا زائرًا في كلية الحاسبات وتقنيات المعلومات بالأكاديمية البحرية للعلوم والتكنولوجيا والنقل البحري – فرع

القاهرة، ومخطِّطًا للبرمجيات بشركة «لوكسور تكنولوجي» الكندية (أُوكْفِيلْ/ أُونْتارْيُو) منذ ٢٠٠٩م إلى الآن، وخبيرًا للغويات الحاسوبية والبرمجيات لمشروع «معجم الدوحة التاريخي» بين عامَيْ ٢٠١٤م و٢٠١٦م.

الدكتور/ مُحَمَّد عبد المنعم عَفِيفي



حصل من جامعة القاهرة على درجة الدكتوراه في هندسة الحاسبات. يَعمل - في الوقت الحالي - مُدِيرًا لأبحاث الصَّوت بمعامِل شركة مايكروسوفت - القاهرة. عملَ باحثًا في معمل Bell وشركة BBN وشركة IBM كما عملَ أستاذًا مُشاركًا بكُلِّيَّة الحاسبات والمعلومات في جامعة القاهِرة. نَشَرَ ما يربو على ٤٠ ورقة بَحثِيَّة في دوريَّاتٍ علميَّةٍ ومُؤتمراتٍ دوليَّةٍ مُتَخَصِّصة.

الدكتور/ شريف مهدي عبده



حصل على درجة الدكتوراه في هندسة الحاسبات عام ٢٠٠٣م من جامعة ميامى بالولايات المتحدة الأمريكية. يعمل حاليا أستاذًا ورئيسًا لقسم تكنولوجيا المعلومات بكُلِّيَّة الحاسبات والمعلومات في جامعة القاهِرة؛ بالإضافة إلى عملِه استشاريًّا لتقنيات مُعالجَة اللُّغة العَرَبِيَّة فى عدد من المراكز البحثية. عملَ - لفترةٍ - باحثًا بشركة BBN الأمريكيَّة، وقد تلقَّى تدريباتٍ عمليَّة في معامل مُعالجَة اللُّغات Bell Labs بشركة «لوسنت» ومركز أبحاث اللُّغات في جامعة كولورادو الأمريكِيَّة. نَشَرَ ما يربو على ٨٠ ورقة بَحثِيَّة في دوريَّاتٍ علميَّةٍ ومُؤتمراتٍ دوليَّةٍ مُتَخَصِّصة؛ كما حصلَ على براءة اختراعٍ عن تقنية (حفص®).



الدكتور/ علي علي فهمي

هو العَمِيدُ السَّابق لكُلِّيَّة الحاسبات والمعلومات في جامعة القاهرة؛ يعمل - في الوقت الحاليّ - أستاذًا في الذَّكاء الاصطناعيّ وتعلم الآلة. عملَ خلال الفترة من ٢٠٠٥ إلى ٢٠١٠ مُديرًا لمركز التميز في التَّنقيب في البيانات ونمذَجة اللُّغة DMCM في مِصر، وله إسهاماتٌ بحثيَّةٌ بارزةٌ في تقنيات اللُّغة العَرَبيَّة وتطبيقاتِها.

مباحث لغوية ٥٨

تطبيقـاتٌ أساسيَّـة فـي المُعالَجة الآليَّة للُّغة العربيَّة

يُصدِر مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية هذا الكتاب ضمن سلسلة (مباحث لغوية)، وذلك وفق خطة عمل مقسمة إلى مراحل، لموضوعات علمية رأى المركز حاجة المكتبة اللغوية العربية إليها، أو إلى بدء النشاط البحثي فيها، واجتهد في استكتاب نخبة من المحررين والمؤلفين للنهوض بعنوانات هذه السلسلة على أكمل وجه.

ويهدف المركز من وراء ذلك إلى تنشيط العمل في المجالات التي تُنَبّه إليها هذه السلسلة، سواء أكان العمل علميا بحثيا، أم عمليا تنفيذيا، ويدعو المركز الباحثين كافة من أنحاء العالم إلى المساهمة في هذه السلسلة.

وتودّ الأمانة العامة أن تشيد بجهد السادة المؤلفين، وجُهد مُحرِّرَي الكتاب، على ما تفضلوا به من رؤى وأفكار لخدمة العربية في هذا السياق البحثي.

والشكر والتقدير الوافر لمعالي وزير التعليم المشرف العام على المركز، الذي يحث على كل ما من شأنه تثبيت الهوية اللغوية العربية، وتمتينها، وفق رؤية استشرافية محققة لتوجيهات قيادتنا الحكيمة.

والدعوة موجّهة إلى جميع المختصين والمهتمين للتواصل مع المركز؛ لبناء المشروعات العلمية، وتكثيف الجهود، والتكامل نحو تمكين لغتنا العربية، وتحقيق وجودها السامي في مجالات الحياة.

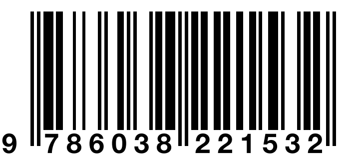
الأمين العام للمركز

أ.د. محمود إسماعيل صالح



ص.ب ١٢٥٠٠ الرياض ١١٤٧٣
هاتف:٠٠٩٦٦١١٢٥٨٧٢٦٨ - ٠٠٩٦٦١١٢٥٨١٠٨٢
البريد الإليكتروني: nashr@kaica.org.sa




9 786038 221532